ERIC Document Resume | See Instruction Sheet | DATE OF RESUME | DATE RECEIVED

A. PUBLICATION DATA

1. ERIC NUMBER: ED001276
2. ORIGINATOR CONTROL NO(S).: NA
3. IDENTIFICATION NO.: CR 1262

DO NOT WRITE IN SPACE BELOW

MICROFILM COPY AVAILABLE: YES ☒ NO ☐

COPYRIGHT GRANTED: YES ☐ NO ☒

COPYRIGHT RELEASE FILED: YES ☐ NO ☒

TYPE OF REPORT:
- PROPOSAL ☐
- IN-PROGRESS ☐
- FINAL ☒
- FOLLOW-UP ☐

OTHER ERIC REFERENCES:

ERIC SATELLITE CODE:

4. AUTHOR(S): Turner, Richard L.

5. TITLE OF DOCUMENT ABSTRACTED: Problem Solving Proficiency Among Elementary School Teachers

6. PUBLICATION TITLE: N A.

7. EDITOR(S): N.A.

8. PLACE OF ORIGIN
ORG. Institute of Educational Research
ST. Indiana University
CITY Bloomington, Indiana
TELE. EXT.
INST/LOC CODE

9. PUBLISHER OR SOURCE: Cooperative Research, Office of Edu., Dept. of HEW, Washington, D. C.

10. VOL.	11. PAGES	12. MO./YR.	13. TOTAL PP.
N.A.	N.A.	64	11

B. INDEXING DATA

14. SUBJECT AREA

15. TYPE OF ENDEAVOR: experimental research

16. INDEXING TERMS (6-12)

	Supervisory ratings	Teaching proficiency
Elementary teachers	Teacher attrition	Teaching reinforcement
Institutional setting	Teacher tasks	
Personal-Social characteristics	Teacher training	

17. ABSTRACT (MAX. 250 WORDS)

Concern lies with the induction of elementary school teachers into their profession. Focus is on the characteristics of beginning teachers with particular attention given to both the role these characteristics play in the beginning teacher's success and how certain of these characteristics are modified as teaching experience increases. It was found, that teachers with very similar characteristics but teaching in different types of settings may find themselves at roughly opposite ends of the success continuum as defined by supervisory appraisals. Which particular teacher characteristics will be associated with success in a given institutional context appears to depend primarily on the socio-economic class composition of the students in this context. The relationship between the institutional context and changes in task performance appear to be only slightly different than those between the former variables and teacher success. Findings force a shift in position from the assumption that statements about teacher skill, success, proficiency or effectiveness are statements about some attributes of a teacher toward the assumptions that such statements are about the relationship between the characteristics of a teacher and the institutional context within which he teaches. Also indicated is a shift to the position that the central re-enforcers in teaching lie in the classroom.

ERIC No.

18. SOURCE OF ABSTRACT

C. FUNDING DATA

19. SUPPORTING ORGANIZATION

20. DIVISION OR DEPT.

21. CONTRACT NO.

PROJECT NO.

22. TOTAL COST OF PROJECT

23. DURATION OR PROJECT

U.S. Department of Health, Education, and Welfare
Office of Education

RIC Form 1

2

L of C

ED001276

PROBLEM SOLVING PROFICIENCY
AMONG
ELEMENTARY SCHOOL TEACHERS

Cooperative Research Project Number 1262

RICHARD L. TURNER

REPRINTED BY PERMISSION
U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN PRINTED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Institute of Educational Research
Indiana University

1276

112/

PROBLEM SOLVING PROFICIENCY AMONG ELEMENTARY SCHOOL TEACHERS

Cooperative Research Project Number 1262

Richard L. Turner

Institute of Educational Research
Indiana University
Bloomington, Indiana

1964

The Research reported herein was supported by the Cooperative Research Program of the Office of Education, U.S. Department of Health, Education and Welfare

Errata

1. The Reference numbers to the bibliography are in error by one following the fourth entry. (6) should be (7), (7) should be (8), etc.
2. Page 3, line 29 threatening and reinforcing.
3. Page 41, Table 25 F and p values appear in row for "Error 1" rather than row for "Tests".
4. Page 69, Table 37, delete "AND TCS TASKS".
5. Page 72, last line, $n = 84$.
6. Page 72a, last two lines, delete p at $.05 = .21$, and $n = 84$.
7. Page 82, line 27, their for there.

ACKNOWLEDGEMENTS

Appreciation is extended to the beginning teachers, the prinipals, supervisors and superinterdents in the 13 Indiana School Systems participating in the project. **Special** gratitude is due the supervisors and superintendents of these systems for their long term commitment to the project and their unflagging support of it.

Grateful acknowledgement is due to Nancy B. Davis and Kinnard White whose work and insights were extremely helpful throughout the project.

Finally, a word of thanks to Mary Anne Bales, Judith Carefoot, Marilynn Papsdorf and Marilyn Badessa whose help with scoring, typing and data processing is greatly appreciated.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ii

LIST OF TABLES iv

LIST OF FIGURES vii

CHAPTER Page

I INTRODUCTION 1

Objectives 4
Instrumentation 5

II OVERVIEW OF PROCEDURES AND SAMPLE CHARACTERISTICS 12

III PROCEDURAL DETAILS AND RESULTS 31

Variables Influencing Performance Change 31
Variables Associated with Supervisory Appraisals of Teaching Success 51
Variables Associated with Teacher Mobility 59
Variables Differentiating Non-Mobile Teachers by Grade Level and System Type 63
Interrelationships Among Task Performance and Personal-Social Variables 67

IV HYPOTHESIS GENERATION AND SYNTHESIS OF RESULTS 79

Hypothesis Generation 79
Synthesis of Results 98

LIST OF TABLES

Table Page

1. Location, Population Base, Pupil Enrollment and Wealth of Participating School Systems 13

2. Teachers Included and Excluded From the Study by System and by Grade Level 16

3. Distribution of Missing Scores by Grade Level and System 17

4. Provision of Supervisor Services in Participating Systems According to System Wealth 21

5. Mean and Range of Number of Quarter Hours of Supervision Given the 1961 and 1962 Samples, by Wealth of System 22

6. Variation in Mean Ratings by Grade Level, and by Assessed Valuation of Systems 24

7. Socio-Economic Class Background of Children in Classroom Units, by Wealth of System 27

8. Means for Pre-Test and Post-Test MTT Scores by System Type and Grade Level 32

9. Analysis of Variance for Pre-Test and Post-Test MTT Scores by System Type and Grade Level 33

10. Means for Pre-Test and Post-Test TTR Scores by System Type and Grade Level 33

11. Analysis of Variance for Pre-Test and Post-Test TTR Scores by System Type and Grade Level 34

12. Means for Pre-Test and Post-Test MTT Scores for Intermediate Teachers in Pattern A Systems by Amount of Supervision Received 35

13. Analysis of Variance of Means in Table 12 35

14. Means for Pre-Test and Post-Test MTT Scores for Intermediate Teachers in Pattern B Systems by Amount of Supervision Received 36

15. Analysis of Variance of Means in Table 14 36

Table Page

16. Means for Pre-Test and Post-Test TTR Scores for Intermediate Teachers in Pattern A Systems by Amount of Supervision Received 37

17. Analysis of Variance of Means in Table 16 37

18. Means for Pre-Test and Post-Test TTR Scores for Intermediate Teachers in Pattern B Systems by Amount of Supervision Received 37

19. Analysis of Variance for Means in Table 18 38

20. Means for Pre-Test and Post-Test MTT Scores for Primary Teachers in Pattern A Systems by Amount of Supervision Received 39

21. Analysis of Variance for Means in Table 20 39

22. Means of Pre-Test and Post-Test MTT Scores for Primary Teachers in Pattern B Systems by Amount of Supervision Received 39

23. Analysis of Variance for Means in Table 22 40

24. Means for Pre-Test and Post-Test TTR Scores for Primary Teachers in Pattern A Systems by Amount of Supervision Received 41

25. Analysis of Variance for Means in Table 24 41

26. Means for Pre-Test and Post-Test TTR Scores for Primary Teachers in Pattern B Systems by Amount of Supervision Received 41

27. Analysis of Variance of Means in Table 26 42

28. Means of Pre-Test and Post-Test MTT Scores for Intermediate Teachers Participating and Not Participating in In-Service Activity Related to Teaching Arithmetic in Pattern B Systems 43

29. Analysis of Variance of Means in Table 28 43

30. Means of Pre-Test and Post-Test TTR Scores for Intermediate Teachers by Hours Per Week Class Schedule in Teaching Reading in Pattern A Systems 45

Table | Page

31. Analysis of Variance of Means in Table 30 45

32. Means of Pre-Test and Post-Test TTR Scores for Intermediate Teachers by Hours Per Week Class Schedule in Teaching Reading in Pattern B Systems 45

33. Analysis of Variance of Means in Table 32 46

34. Quarter Hours of Supervision Received by Groups Categorized by Types of Difficulties in Teaching, by System Types 50

35. Number of Teachers in Pattern A Systems Appraised as Above Average-Superior or as Average-Below Average, by Type of Teaching Difficulty 52

35a. Number of Teachers in Pattern B Systems Appraised as Above Average-Superior or as Average-Below Average by Type of Teaching Difficulty 53

36. Correlations Among Change Scores on MTT and TTR Tasks, and TCS Scale Scores for Intermediate Teachers 68

37. Correlations Among Change Scores on MTT and TTR Tasks, and TCS Scale Scores for Primary Teachers 69

38. Intercorrelations Among MTT Pre-Test, TTR Pre-Test and TCS Variables for Intermediate Teachers 71

39. Intercorrelations Among MTT Pre-Test, TTR Pre-Test and TCS Variables for Primary Teachers 72

40. Factors and Factor Loadings Associated with the MTT, TTR and TCS for Primary and Intermediate Teachers 72a

41. Factors and Factor Loadings Associated with MTT Tasks, TTR Tasks and TCS Scales for Primary and Intermediate Teachers 74

LIST OF FIGURES

Figure Page

1. Changes in MTT and TTR Scores by Months of Teaching Experience 47

2. Differences in Mean Z Scores on MTT, TTR, and TCS Factors for "high" "low" Groups, Pattern A Systems 54

3. Differences in Mean Z Scores on MTT, TTR, and TCS Factors for "high" and "low" Groups, Pattern B Systems 55

4. Differences in Mean Z Scores on TCS Factors for "high" and "low" Groups Teaching in Working Class or Middle Class Schools in Pattern A Systems 57

5. Differences in Mean Z Scores on TCS Factors for "high" and "low" Groups Teaching in Working Class and Middle Class Schools in Pattern B Systems 58

6. Mean Z Scores on MTT, TTR and TCS Factors for Mobile and Non-Mobile Intermediate Teachers by System Type 61

7. Mean Z Scores on MTT, TTR and TCS Factors for Mobile and Non-Mobile Primary Teachers by System Type 62

8. Mean Z Scores for MTT, TTR and TCS Factors for Non-Mobile Teachers by Grade Level and System Type 64

9. Mean Z Scores for Non-Mobile Teachers Appraised as Average-Below Average or Above Average-Superior, by System Type 66

CHAPTER I

INTRODUCTION

This study is broadly concerned with the induction of elementary school teachers into their profession. The focus of the study is on the characteristics of beginning teachers, with particular attention to both the role these characteristics play in beginning teacher success and how certain of these characteristics are modified as teaching experience increases. Underlying the study is a particular perspective of teaching as an occupation. Teaching is seen not only as a type of work which shares many common features with other types of work, but also as having certain unique features.

As is true of other occupations, teaching involves the performance of a body of work tasks. The work tasks performed by teachers provide a basis for distinguishing this occupation from others and within the occupation, for distinguishing occupational subgroups. As is also true in other occupations, the worker or practitioner brings his own personal characteristics to his work tasks. These personal characteristics, especially those of a personal-social kind, may be viewed as providing a background, indeed, a personally generated context, within which the work tasks are performed. The importance of this personally generated context undoubtedly varies from occupation to occupation, but that it is important in teaching and the other professions can hardly be doubted. Finally, teaching, and especially public school teaching, like a great many other occupations in an industrial society, occurs in an institutional context. By its very definition, public school teaching is inseparable from its institutional context viz., the public school system.*

Those features of teaching which make it unique, that is, which differentiate it from the managerial occupations as well as other professions, lie in its work tasks. Not all of these tasks, however, are unique to it. Among the important work tasks for teachers is the setting of tasks for pupils. Teachers specifically determine how pupils are to occupy themselves. The types of tasks with which teachers occupy pupils are diverse. They vary with the grade level taught by the teacher or, more generally, with the level of development of the pupil, and also with the goals of instruction. The task may be for a fourth grade pupil to solve a group of multiplication examples, or, in a very permissive context, the teacher may set a pupil the task of determining how he should occupy himself in learning arithmetic.

---

*The relationships between teaching and other occupations in this paragraph were generated on the basis of Edward Gross's, Work and Society. (6)

Pupil task setting, however, is not a work task unique to teaching. Managerial personnel, for example, have a broad responsibility for determining how workers are to occupy their time. Indeed the more business-like the organization, the more specificity in how the worker is to occupy homself. The parallel between this work task of the teacher and certain work tasks of managers, then, is close. Task setting is not unique to teaching.

A second type of work task for teachers is the appraisal of pupil behavior in relation to the tasks set for the pupil. This set of work tasks presumes the existence of standards or criteria against which pupil responses are judged. These criteria may be held by the teacher at different levels of complexity or elaborateness, and he may hold them with different degrees of flexibility or rigidity. For instance, one teacher may hold rapid, accurate calculation in arithmetic as a relatively simple criterion for students to achieve, but he may hold this criterion very rigidly, i.e. all students must meet an absolute standard in terms of speed and accuracy. Another teacher might hold much more elaborate criteria in arithmetic, involving perhaps speed and accuracy, but also involving extended comprehension of number systems, the associative, distributive and commutative principles, and ability to apply numerical concepts in many types of situations, yet such a set of criteria might be held flexibly, so that the degree to which a child is expected to attain any part of the set is adjusted to the ability level of the child. In any event, the appraisal of pupil responses against a criterion or set of criteria is not unique to teaching. Again, managerial personnel regularly appraise the responses of workers against an established criterion, frequently as a part of the over-all quality control functions of the organization. Indeed, one might view this work task of the teacher as a "quality control" task.

The third among the work tasks of teachers exists primarily as a consequence of performance of the other two work tasks, and it is this work task that seems unique to what we call "teaching", whether it be in the public school or elsewhere. A discrepancy between a pupil's actual responses to a task set for him by a teacher and those responses required by the criterion or standard held by the teacher generates a specific instructional problem. The problem is to bring the observed responses of the pupil in line with the responses held as criterial by the teacher. In order to resolve specific instructional problems, the teacher may be viewed as engaging in extensive instrumental responding, responding which has as its objective closing the gap between the observed responses of the pupil and the criterial responses held by the teacher. The skillfulness with which a teacher can close these gaps between what a pupil can actually do and what the teacher expects him to be able to do is an important component of skill in teaching. A component which might be labeled the "instructional work task" of the teacher. It is important to recognize, however, that the instructional work task of the teacher is only one of three components of teaching, and that the exact nature of specific instructional tasks is contingent on the criteria or standards held by the teacher in relation to the type of pupil work tasks set. A teacher who holds elaborate criteria potentially has much more difficult instructional problems than a teacher who holds simple criteria, given the same level of pupil ability.

In summary, there are three main work tasks in teaching: the setting of tasks for pupils, the appraisal of pupil responses to these tasks against criterial responses, and, as a consequence of discrepancy between observed and criterial responses to a task, an instructional work task, involving extensive instrumental responding on the part of the teacher to close the gap between observed and expected pupil responses.

In the perspective underlying this paper, the teacher's performance of his work tasks is viewed as open to modification. Modification of work task performance is further seen as ultimately traceable to the effects of threats of deprivation, reinforcers and social sanctions operating in the environments in which teachers teach. At present, however, the loci of these threats of deprivation reinforcers and sanctions has not been clearly identified, hence one concern of the present study was to tentatively locate these loci or sources of stimuli which seem to influence or control teacher task performance. That the principal sources of threatening reinforcing stimuli lie somewhere in the institutional context in which the teacher teaches, is, of course, assumed.

That changes in teacher task performance may be modified through threatening or reinforcing stimuli issuing from one or another aspect of the institutional context, does not abrogate the possibility that underlying task performance are certain relatively stable abilities and, possibly, certain personal-social characteristics, which hinder or facilitate changes in performance. Indeed, one might reasonably hypothesize that this is the case and therefore conduct a search for such abilities and characteristics. This search has been conducted in the studies (11, 12, 14) preceding this one and was continued in the present study.

Unlike the performance of the work tasks of teaching, the personal-social characteristics which enter into the personal context generated by the teacher may be viewed as relatively stable. While these characteristics are probably acquired by each individual teacher, their acquisition may be viewed as appraoching completion by the time the teacher enters the profession. Within the present conceptualization, then, the personal-social characteristics through which the personal context aspect of teaching is generated are viewed as distinctly less modifiable by either teacher preparation or teaching experience than is task performance. The problem, therefore, is not to determine how such personal characteristics change, but rather to determine the relationships these characteristics hold to task performance.

The important point is that one may distinguish between the work tasks of a teacher and the teacher as a person. This distinction subsequently permits the problem to emerge as one of finding out how the personal characteristics of the teacher relate to the performance of his work tasks. Consequently, one of the problems with which the present study dealt was whether the relationship between the personal characteristics of the teacher and his performance of his work tasks is a direct relationship or whether it is more nearly an indirect or "interactional" relationship.

Finally, it is intrinsic to the theoretical viewpoint underlying this paper that, as an occupation, teaching goes on in a definite environment or institutional context. While both the work tasks of the teacher and the personally generated context within which these work tasks are performed may be distinguished from the institutional context, they clearly cannot be treated as independent of it. The authority structure in the school system in which a teacher teaches controls many of the sanctions teachers receive and therefore has an effective means of controlling the behavior of the teacher. Moreover, the authority structure of the school system serves as a mediator of the broader sanctions held by the community. The persons in this structure interpret, as it were, what the community requires by way of education for the young, and, in this sense, hold the sanctions of the larger community of which they are a part.

While one may expect, then, that the institutional context in which teaching occurs will affect that teaching, which particular factors or variables in this context are influential is not well understood. Thus a major problem in the present study was to identify and determine the influences of selected variables operating in the institutional context.

Objectives

In accord with the preceding discussion, the variables observed and related together in the present study may be viewed as belonging to three main sets, (1) those bearing on work task performance, (2) personal-social variables bearing on the personally generated context of teaching, and (3) variables associated with the institutional context in which teaching occurs. The objectives of the study may consequently be cast in terms of these three sets of variables.

One objective of the study was to determine how task performance in teaching arithmetic and in teaching reading are associated with each other, and how each of these types of performance relate to the various personal-social factors assessed by the Teacher Characteristics Schedule developed by D.G. Ryans (7).

A second objective was to determine how task performance in teaching arithmetic and reading changed during the first two years of teaching experience. Especially important in this respect was observation of how the personal-social characteristics of the teacher, grade level taught, and selected variables issuing from the institutional context of teaching, including school system wealth, amount of supervision given, organization of supervisory services, and ratio of working class to middle class pupils, amount of in-service education, and number of hours per week class schedules influenced changes in teacher task performance as experience increased.

A third objective of the study was to ascertain how task performance and personal-social factors are related to various aspects of teaching success, including supervisory appraisals of success, teaching difficulties or problems experienced by beginning teachers, and amount of supervision "needed".

A fourth objective was to determine the relationship of task performance variables and personal-social variables to teacher mobility and attrition during the first two years of teaching experience in different types of institutional settings.

These four objectives were used primarily to structure the study and point out which particular relationships among which particular variables were to be examined. Underlying them, however, was a much more general question concerning whether it is profitable to talk about teaching and teacher competence independent of the specific type of institutional context in which the teacher practices his profession. This question clearly cannot be answered in any single research, however in the present instance it served as a nucleus around which many of the procedures, analyses of data and interpretations of results were organized.

Instrumentation

To achieve the objectives set forth above, three sets of instruments were used: the first set dealt with task performance in teaching arithmetic and in teaching reading; the second set dealt with the personal-social characteristics of teachers; the third, which involved relatively unrefined instrumentation, with the practices, observations and appraisals of elementary supervisors and principals in relation to their beginning teachers.

Task Performance Instruments. Two paper and pencil task performance instruments were used: Intermediate Grades Mathematics Teaching Tasks, Form E_1 (hereafter called MTT)(10) and Elementary Grades Teaching Tasks in Reading, Form B (hereafter called TTR) (2).

Form E_1 of the MTT consists of four separate tasks. Task 1 is focused on the middle grades (3 and 4) and requires the teacher to relate a set of arithmetic exercises to each of ten specific objectives varying in degree of relevance to the exercises. The relevance of each objective to the exercises is rated by the teacher on a four-point scale; a maximum of ten points, one for each objective correctly scaled, may be obtained on this task by the respondent. Performance time on this task is typically about 10 minutes. Task 2 is centered in grade 5, and requires the teacher to score a group of ten two-digit divisor division examples which have been solved by a fifth grade pupil. Four of these examples are erroneous. Two of the four are random (non-replicated) errors in their class while two are systematic (replicated errors) in their class. After scoring, the teacher must first decide whether to interview the student about his errors, or whether not to do so. A decision to interview the student subsequently requires the teacher to rate each of ten statements of the nature of the errors made by the student on a three-point scale. A respondent who decides not to interview the student receives four points for this decision which is approximately equal to chance responding to the ten statements,* while a respondent who decides to interview receives one point for each statement correctly scaled, with a maximum of ten points possible. Performance time on this task is typically about 12 minutes.

Task 3 is identical in structure to Task 2, but it is focused in grade four and involves different content than Task 2. The content of Task 3 is diverse, covering the four basic arithmetic operations at a level of attainment typical in grade four, plus understanding of large numbers and fractions. This sort of content is frequently covered by the term "mixed arithmetic exercises" and may be used by teachers for preliminary diagnosis of pupil difficulties. The teacher responds to the replicated and non-replicated errors as she sees them in the same way as in Task 2, and the responses are scored on the same basis. Performance time is typically about 12 minutes. Task 4 is one which may be encountered in any of the intermediate grades. It requires the teacher to rank seven division examples according to their difficulty level for pupils as established empirically. This task is scored by counting one point for each example correctly ranked and one-half point each for reversed adjacent ranks. The total score received is doubled in this task, hence a maximum of 14 points is possible. Performance time on the task is about five minutes.

---

*Failure to give four points for this decision reduces the reliability of this task, since some respondents decide to interview but respond to the ten statements on a basis not distinguishable from random responding.

Among the three major work tasks in teaching, as described earlier, these four arithmetic teaching tasks focus primarily on the "pupil task selection" aspect of that teacher work task involving setting tasks for pupils, exemplified by Tasks 1 and 4, and on comparing actual pupil responses to criterial responses, as exemplified by Tasks 2 and 3. Hence, these four tasks do not sample the "instructional" work tasks of the teacher. From a research viewpoint this fact is important because it marks a fundamental difference between the work of the author and the work of Smith and Meux (9). The latter have been primarily concerned with categorizing the logical behavior of the teacher as he performs the instructional work task in the classroom, while the writer has been concerned with the measurement of performance on work tasks precursory to the instructional work task.

The validity of the MTT has been examined in a number of studies previous to the present one (11, 12 and 13). The most germane results of the earliest among these studies are as follows: First, scores on the MTT significantly increased among preparatory elementary teachers between the time they entered an arithmetic methods course and the completion of student teaching nine months later. Within this group, student teachers of grades 4-6 showed significantly greater increases than student teachers of grades 1-3. Second, teachers of grades 4-6 with 1-10 years experience performed significantly higher than did student teachers of grades 4-6. The former group also performed significantly higher than experienced teachers of grade 3, drawn from the same school systems. Third, among experienced teachers, those graduated from large public institutions performed significantly higher than those graduated from small, private institutions, and those teaching in school systems with population bases of greater than 25,000 performed better than those teachers in systems with population bases of less than 25,000. Fourth, a follow-up study by White (13) of the group measured prior to their methods course and after the completion student teaching indicated that those who had low MTT scores at the end of student teaching were more likely ($p < .10$) to leave their first position at the end of the first year than those who had high MTT scores.

In toto these results suggested that performance on the MTT increased with teacher preparation and experience, which supported the principal hypothesis in the theoretical perspective utilized. The study by White, however, also indicated that selection may have played a role in the performance differences between student teachers and experienced teachers. The present study was in part designed to help disentangle the effects of learning from the effects of selection.

The latest of the validity studies (13) with the MTT was undertaken after the present study was begun. In the study, all teachers, grades 3-6, in a single school system were assessed, and the resulting scores related to supervisor ratings of skill in teaching arithmetic and, among teachers of grades 4 and 5, to which the tasks are most germane, to pupil achievement in arithmetic. The results indicated that teachers who were rated high in skill in teaching arithmetic had significantly higher MTT performance than those rated low, and that pupils who passed between pairs of high scoring teachers in grades 4 and 5 were significantly higher in arithmetic achievement by the middle of grade 5, when previous achievement and intelligence were controlled, than were the pupils who passed between pairs of low scoring teachers in these grades. The latter study, together with the earlier ones, distinctly suggest that the MTT is a relatively valid measure of skill in teaching arithmetic in the intermediate grades. There is, however, no evidence that it is a valid measure of skill in teaching arithmetic in the primary grades.

The reliability of the MTT as well as its validity was an important consideration in the present study. To obtain the reliability of this instrument, an alternate for each task appearing in Form E_1 was developed. These alternate tasks were then collected in an alternate form of the instrument, Form F_1. The two forms were administered to a group of 63 preparatory elementary education students in a single sitting. Part of the group took Form E_1 first and Form F_1 second, and part took Form F_1 first and Form E_1 second. The results indicated that the mean increases in score between forms was less than one point, while coefficient of equivalence between forms was .754. Subsequent examination of the equivalence of each task to its alternate indicate that Task 2 in Form F_1 correlated only .05 with Task 2 in From E_1; hence, these two tasks were not at all parallel and served to reduce the over-all parallel form reliability of the two instruments. Because of the failure of Task 2 in Form F_1 to function an an equivalent of Task 2 in Form E_1, and because the validity estimates were based wholly on Form E_1, Form F was not utilized in the present study. While these results were disappointing from the viewpoint of parallel form reliability, they were encouraging from the viewpoint of practice effects, since taking one form increased performance on the second form by less than one unit.

Form B of Teaching Tasks in Reading also consists of four tasks. Task 1 is focused in the primary grades and requires the teacher to categorize and label ten errors in word recognition made by a pupil. The maximum possible score is ten, and the task requires about ten minutes performance time. Task 2 is focused primarily on the middle grades (2 to 5) and requires the teacher to rank seven paragraphs in order of their reading difficulty level for pupils as determined by readability formulas. The maximum score of this task is six and performance time is about ten minutes. Task 3 centers in the intermediate grades and requires the teacher to categorize and label ten syllabication errors made by a pupil. Only categorization is scored in this task. The maximum score is ten, and about ten minutes is required to perform the task. Task 4 focuses in the upper primary grades and requires the teacher to match from a pool of fifteen exercises, six pairs of exercises, each pair of which builds the same reading skills. The maximum score is eight, and performance time is about ten minutes.

As in the MTT, the four tasks in the TTR deal primarily with pupil task setting (Tasks 2 and 4) and with the comparison of pupil responses to criterial responses (Tasks 1 and 3). In addition, the operations required of the teacher on the tasks in the TTR are similar to those required on the MTT. Tasks 1 and 3 in the TTR and Tasks 2 and 3 in the MTT all require the teacher to categorize pupil errors, although in the TTR, teachers are not asked to judge the significance of these errors. The ability underlying these categorization operations may be generally conceptualized as concept formation ability. Tasks 2 and 4 on the TTR and Tasks 1 and 4 on the MTT, more nearly require the utilization of previously acquired concepts as these apply to the selection and organization of learning materials although concept formation abilities are probably also involved. For instance, in Task 2 on the TTR and Task 4 on the MTT, the teacher must have a knowledge of those variables which increase the difficulty level of reading passages and of division examples, respectively, in order to correctly categorize the passages and examples and complete the tasks. All told, the similarities of the MTT and TTR are similarities in cognitive operations, while the differences between these two are centrally differences in content; i.e., arithmetic and reading.

The validity of Form B was not established when the present study began. However, the tasks in Form B were adapted by N.B. Davis from those which best discriminated between experienced and preparatory teachers in an earlier study by E.W. Wade (14). Subsequently, Davis showed that performance on the tasks in Form B significantly increased among experienced teachers taking an in service methods course in reading, and Bradtmueller (3) showed that performance on these tasks significantly increased both with a methods course in language arts and with student teaching experience.

These findings conformed to the general hypothesis that performance changes as a function of teacher preparation and experience, but beyond this there was no evidence of validity.

To determine the reliability of Form B, four alternate tasks were developed by N.B. Davis during the project. These tasks were collected into an alternate form, Form C. These two forms were subsequently administered at a single sitting to a group of 41 graduate students enrolled in a course in the teaching of reading. Part of this group took Form B first and Form C second, and part took Form C First and Form B second. The results indicated that there were no practice effects between forms, while the correlation between the two forms was .784. Although the parallel form reliability coefficient of this magnitude is reasonably good, Form C was not subsequently employed in the present study. To employ it would have introduced an additional variable in the study, and the study included too many other variables to warrant introduction of another via the use of alternate forms of the TTR.

Instruments Measuring Personal-Social Characteristics. Form E 54 of the Teacher Characteristics Schedule, developed by D.G. Ryans (7), was utilized to assess personal-social characteristics of teachers relevant to classroom teaching. This instrument contains ten "factors" or scales. Of these factors, three factors, X_{co}, Y_{co} and Z_{co}, are bi-polar, correlate ($_{c}o$) scales which hold correlate status with observed patterns of elementary teacher classroom behaviors. In addition, two of the scales, Z_{co} and Y_{co}, hold relationships to patterns of pupil behavior (8, pp. 97-103 and 250-51. Factor Z_{co} reflects stimulating, imaginative, original vs. dull, routine teacher classroom behavior, and is related to the teacher's ability to encourage pupil participation and initiative. Factor Y_{co} reflects business-like, systematic, and responsible vs. unorganized and slipshod classroom behavior and is related to constructive, responsible, cooperative and controlled pupil activity. Factor X_{co} reflects understanding, friendly, democratic vs. aloof, harsh and autocratic teacher classroom behavior. According to Ryan's results, however, this scale was not directly related to the pupil classroom behaviors observed in his study.

The remaining seven scales of the TCS encompass the following characteristics: R_{co} (favorable vs. unfavorable attitudes toward pupils); $R1_{co}$ (favorable vs. unfavorable attitudes toward democratic pupil practices); Q_{co} (favorable vs. unfavorable attitudes toward school staff personnel, especially teachers); B_{co} (traditional, subject-centered viewpoint vs. child-centered, permissive viewpoint); I_{co} (verbal understanding); S_{co} (emotional adjustment) and V_{co} (tendency to use self-enhancing and socially acceptable responses)*.

---

*V_{co} was not employed in the present study. This scale has low reliability and its validity has not been well established.

These scales were not directly related to the observed classroom behavior of teachers.

In addition to examining the relationships between the TCS correlate scales and observed classroom behavior, Ryans also examined them in relation to principal's ratings. Using scoring Key 111, which is that used in the present study, he found that elementary teachers rated as "superior" in over-all performance were significantly higher than those rated "poor" in over-all performance, on the following scales: X_{co}, Y_{co}, Z_{co}, R_{co}, B_{co}, I_{co} and S_{co}. These findings were based, however, on only 60 teachers (35 "superior", 25 "poor") from an initial nation-wide mail sample of 200 teachers; hence, caution must be exercised in their interpretation. (8. 261.).

The split half reliability of the TCS scales among elementary teachers ranges from a low of .64 for V_{co} to a high of .89 for X_{co}. About half of the remaining scales have reliability coefficients in the seventies, and about half in the eighties. These were interpreted as satisfactory. While no data on the stability of the scales for elementary teachers are provided by Ryans, he does provide these data for secondary teachers. With a one year interval between testings, the stability coefficients for 112 secondary teachers were as follows: $X_{co}=.73$, $Y_{co}=.63$, $Z_{co}=.34$, $R_{co}=.77$, $R_{1co}=.70$, $Q_{co}=.68$, $B_{co}=.73$, $I_{co}=.78$, $S_{co}=.22$, $V_{co}=.22$. Of the scales, Z_{co}, S_{co} and V_{co} are apparently unstable, although the remaining stability coefficients seem satisfactory. Whether the stability of these factors is similar for elementary teachers is an unanswered question; however, there is no apparent reason to hypothesize that personal-social variables of this type would be less stable among elementary than among secondary teachers.

The MTT, the TTR, and the TCS were the only instruments utilized directly with beginning teachers in the study. Information concerning supervisory practices and appraisals and other data about schools and school systems were obtained via questionnaire and interview, and are more closely related to procedural matters than to instrumentation in the strict sense. Hence, the manner in which such information was obtained is noted, as appropriate, in subsequent sections.

CHAPTER II

OVERVIEW OF PROCEDURES AND SAMPLE CHARACTERISTICS

In the spring of 1961, 25 Indiana school systems were invited to participate in the study. "Participation" entailed the following commitments: (1) the system would remain in the study for two years; (2) it would permit the beginning, inexperienced teachers, grades 1-6, employed in 1961, to be tested in the fall of 1961 and re-tested in the spring of 1963 and the beginning inexperienced teachers employed in 1962 to be tested in the fall of 1962 and re-tested in the spring of 1963; (3) it would cooperate in supplying information about the system and about the performance of the beginning teachers. Commitment to these conditions represented a substantial investment on the part of a system since about six hours of beginning teacher time and several hours of the time of supervisors and principals would be involved.

Of the 25 systems invited, two declined participation because of other research commitments, ten agreed to participate but were later excluded because they did not employ beginning, inexperienced teachers in 1961, while 13 systems both agreed to participate and employed beginning, inexperienced teachers in 1961. These 13 participating systems are not a random sample of Indiana school systems. They are, however, quite representative of certain types of Indiana systems, principally: (1) suburban systems; (2) systems in highly industrialized communities centering in cities of 25,000-100,000 inhabitants; (3) systems centering in cities of 5,000-25,000 inhabitants. The latter cities are typical midwestern "county seat" communities having some industry, but also serving adjacent rural areas. The 13 systems are spread throughout the state, but with somewhat greater concentration in the broad central section which contains the greater proportion of the medium-sized cities in the State. The wealth of the participating systems, as indexed by the average adjusted assessed valuation per resident pupil in average daily attendance, was also reasonably well spread over the total range of school system wealth in the State. The mean wealth of the participating systems, however, was more nearly comparable to that of systems with similar pupil enrollment than to the mean wealth of systems in the State as a whole. The mean wealth of the systems in the sample as of 1960-61 was $7,854 per resident pupil ADA, the mean wealth of the 47 other systems with enrollments similar to the participating systems was $8,187 per resident pupil ADA, while the mean wealth of all systems in the state in 1960-61 was $8,887*. The error of the mean of the participating sample from the mean of systems in the same enrollment categories is predominantly attributable to the failure to include in the sample a relatively wealthy system ($10,000 -$11,000 per resident pupil ADA) with a population base in the 5,000-15,000 class, and enrollment in the 2,000-3,000 category. Two such systems were approached, but did not employ beginning teachers in 1961 and thus were not included. A summary of geographical locations, population base of the systems, enrollment in grades K-12 and wealth for each participating system is shown in Table 1.

---

*Data in this paragraph is based on the following sources: School Board Bulletin for Indiana School Boards, No. 25, September, 26, 1961, and Indiana State Teachers Association Circular No. 9, June, 1961.

TABLE 1. LOCATION, POPULATION BASE, PUPIL ENROLLMENT AND WEALTH OF PARTICIPATING SCHOOL SYSTEMS.

Participating Systems	Location in Indiana	Approximate Population 1960[#]	Enrollment 1-12 1960-61	"Per Pupil Valuation" 1960-61
A	West Central	108,000	19,928	$ 9,827
B	Central	98,000	9,561	7,424
C	South	51,000	9,837	6,868
D	East Central	49,000	9,480	8,497
E	North	40,000	8,695	8,808
F	Central	34,000	5,111	7,044
G	South Central	30,000	6,642	10,489
H	East Central	20,000	5,621	6,356
I	North Central	15,000	3,554	6,393
J	South Central	15,000	4,136	9,034
K	South Central	14,000	2,943	8,344
L	South Central	13,000	2,641	6,733
M	North Central	10,000	2,410	5,781

[#] Population figures are for the school corporation, not the incorporated city in which the school corporation is centered.

In the late summer and early autumn of 1961, all beginning inexperienced teachers, grades 1-6, in each participating system were invited to participate in the project. The beginning teachers in each system were subsequently tested by the writer on the MTT and the TTR at a common location in the system. At the conclusion of testing, the TCS was handed to each participant to be returned by mail when completed. The earliest testing date was August 28, 1961, and the latest testing date was October 15, 1961. Some teachers therefore had as much as six weeks' classroom experience prior to testing. The testing dates were irregularly distributed over systems, however, so that no systematic confounding of experience before testing with city size or wealth occurred.

In the late summer and fall of 1962, these procedures were repeated with the beginning, inexperienced teachers, grades 1-6, employed for the 1962 school year. In 1962, however, system M as shown in Table 1 did not employ any beginning teachers and was therefore excluded from the 1962 sample. In addition, a few teachers who had graduated from Indiana University the previous June had already been tested on either the MTT or the TTR or both as part of another study. These teachers were not re-tested on the instrument they had completed earlier, and the earlier score was utilized. Again, these few teachers were irregularly distributed over participating systems and no systematic effect in the data occurred.

In the spring of 1963, all teachers who had previously been tested and who were still in the system in which they began were asked to participate once again. Thus the group which began in 1961 was re-tested at the end of two years experience, and the group which began in 1962, at the end of one year of experience. These teachers were tested on the MTT and the TTR, at a common location in their school systems, between April 17 and May 20, 1963. In addition, these teachers completed questions relative to their participation in in-service activity, including workshops and college courses,at this sitting.

The samples of teachers produced through the initial sampling procedures provided the basic pool of teachers for the study. In order to maintain control over several variables, however, a number of teachers were excluded from the main body of the study. First, all teachers who did not teach either reading or arithmetic in one of the grades one to six were excluded.* Second, all teachers who taught only special classes of retarded students were excluded. Third, all teachers who had teaching experience at any level prior to beginning to teach in the elementary school were excluded. Fourth, all teachers who received their Bachelors degree more than one year prior to testing were excluded. The latter exclusion was made in order to maintain homogeneity in the interval between the time the methods courses and student teaching were taken and the time of entry into the profession. These exclusions resulted in a final sample of recently prepared elementary teachers who were directly engaged in teaching reading and/or arithmetic in grades 1-6 and who had no previous full-time teaching experience.

---

*The eight teachers in this group were employed as a control group in a portion of the study.

In addition to the exclusions cited above, two other factors affected the number of teachers involved in certain aspects of the study. First, approximately 20 per cent of the beginning teachers failed to return the TCS.* Second, three teachers who had graduated from Indiana University were listed as members of the sample tested on the MTT or the TTR at the completion of student teaching were found to have been absent for either the MTT or the TTR, but not both, while an additional group of beginners either arrived late or were forced by external circumstances to leave early during the testing session in their school system, and therefore completed only one of the two tests to be taken. Hence, either the MTT or the TTR was missing for the members of these two groups. A summary of the numbers of teachers included and excluded from the study by grade level and by school system is shown in Table 2, while Table 3 shows the distribution, for teachers included in the study, of missing scores on the TCS, MTT, or TTR by grade level and by system.

After the initial testing, participating teachers received no further communications from the investigator until invited to participate once again in the spring of 1963. The invitational letter was placed in the hands of teachers approximately one month prior to testing. Since participating teachers had not been told they would be re-tested until they received the invitational letter, the group beginning in 1961 had about 15 months of classroom experience without knowledge that they would be retested, while the group beginning in 1962 had about 7 months of classroom experience without such knowledge. Each group had approximately one month of classroom experience with knowledge that they would be re-tested, but this month was in all cases at least seven months removed from the initial testing, so that memory of the tests and corresponding self-practice on tasks of the type contained in them were assumed to be minimal.

---

*The MTT and TTR scores of the group failing to return the TCS differed only randomly from the group returning the TCS.

TABLE 2. TEACHERS INCLUDED AND EXCLUDED FROM THE STUDY BY SYSTEM AND BY GRADE LEVEL.

System	Grades 4-6		Grades 1-3		
	Included	Excluded	Included	Excluded	Totals
A	14	1	14	0	29
B	10	1	11	5	27
C	7	0	20	0	27
D	1	1	13	7	22
E	9	1	8	5	23
F	5	0	15	0	20
G	8	0	10	0	18
H	3	0	4	1	8
I	4	0	4	0	8
J	3	1	5	0	9
K	2	0	6	0	8
L	3	1	5	0	9
M	1	2	0	0	3
Totals	70	8	115	18	211

TABLE 3. DISTRIBUTION OF MISSING SCORES BY GRADE LEVEL AND SYSTEM

System	Grades 4-6				Grades 1-3			
		Missing Scores				Missing Scores		
	n	TCS	MTT	TTR	n	TCS	MTT	TTR
A	14	1	1		14	2	1	
B	10	2			11	4		1
C	7	3		1	20	6		
D	1				13	4		1
E	9	1			8			
F	5				15	2	1	1
G	8	3		1	10	2		1
H	3				4			
I	4				4	1		
J	3	1			5	2		1
K	2				6		1	
L	3	1			5	2		
M	1				0			
Totals	70	12	1	2	115	25	3	5
Percent Missing		17.1	1.4	2.9		21.7	2.6	4.3

Superintendents, supervisors and principals in participating schools received communications, including requests for information about teachers, at various times between the initial and final testing of teachers. Both superintendents and supervisors knew from the beginning of the project that beginning teachers in their systems would be retested; principals, however, did not receive this information from the project center. The first direct communication between principals and the investigator occurred at the end of the first year of experience for the 1961 groups; i.e., June, 1962, at which time information about teachers was requested from the "person who had closest supervisory contact with the beginning teacher". This request for information had not been planned when the project was undertaken; rather, the plan was to acquire information on teachers at the time the final testing of teachers was done. Informal communications with superintendents and supervisors indicated, however, that the 1961 sample was eroding at a very high rate and that accurate information about those teachers who were leaving would be lost unless it were obtained immediately. Those persons having closest supervisory contact with beginning teachers were therefore asked to supply the following items of information about their beginning teachers as of June, 1962, for the 1961 sample:

1. Whether the teacher would be in their system the following year, and, if not, the reason for the teacher's withdrawal.
2. Number of pupils in the teacher's class, 1961-62.
3. Number of teachers in the building.
4. Number of hours per week the teacher devoted to teaching reading and to teaching arithmetic.
5. Types of in-service activities in which the teacher participated and number of hours devoted to each type.

In addition, the person having closest supervisory contact was asked to appraise each beginning teacher on a three point-scale upper 25%, middle 50% and lower 25% (high, middle, and low) in each of five areas of teaching: 1. personal-social relationships with other teachers; 2. the perception of other teachers in the building of the beginner's ability to do classroom teaching; 3. the skill of the teacher in teaching arithmetic compared to other teachers in the system; 4. the skill of the teacher in teaching reading compared to other teachers in the system; 5. the over-all skill of the teacher compared to other beginning teachers with whom the respondent had had supervisory contact during the past five years.

The instruction to systems that the person having closest supervisory contact with the beginning teachers in each system was to supply the information resulted in supervisors supplying the requested information in two systems, A and G, while principals supplied the information in the remaining systems.

In May-June of 1963, information on both the 1961 and 1962 samples of beginning teachers was again requested of systems; as in June, 1962, the person having closest supervisory contact with the beginning teacher was requested to supply the same information as supplied in June, 1962, but relative to the 1962-63 school year. Two additional items of information were requested, however: 1. the average number of supervisory contacts per month with each beginning teacher under his direct supervision, and the average time per contact; and 2. a statement concerning the areas of teaching in which the teacher most needed to improve during his first or second year of experience. In June, 1963, all systems interpreted the person who had closest supervisory contact to be the principal, with the exception of system A, in which the supervisors supplied the required information.

In addition to obtaining the information detailed above, primarily from principals, either supervisors or superintendents were interviewed on the day teachers were given the final tests in the spring of 1963. In this interview, the respondent was requested to respond to four items for each participating teacher who was still in his school system.

"1. Which of the following most nearly characterized the type of population from which this teacher's school draws its children?

(a) Lower class (Parents are predominantly unskilled or semi-skilled)

(b) Lower and middle middle class (Parents are skilled or white-collar or sales)

(c) Upper middle class (Parents are professional or managerial.)

(Respondents were encouraged by the interviewer to elaborate on the class background of each school and these elaborations were noted.)

2. On the average, about how frequently have you or another supervisor had supervisory contacts (conference, observation, etc.) with this teacher?

3. What kinds of teaching problems do you feel this teacher performs well, and what kinds do you feel she performs poorly?

4. On the whole, how would you appraise this teacher's overall teaching performance?"

Whether the supervisor or superintendent was interviewed with these questions was contingent on the way in which supervision was conducted in the system. In systems I, K, L and M there were no supervisors, hence superintendents were interviewed. In systems A, D, E, F, G, and H, supervisors were interviewed, while in system B, the Assistant Superintendent (in charge of instruction) was interviewed, and in system C the Director of Elementary Education was interviewed. In all systems except systems A and C all questions in the interview were responded to by the person interviewed. In system A the supervisor interviewed was the same as the person who completed the forms which went to the "person who had closest supervisory contact with the beginning teachers in the system"; hence, this person did not respond to interview questions 3 and 4 since this information was supplied elsewhere. In system C, the person interviewed did not respond to questions 3 and 4 because the direct supervisory role was placed wholly in the hands of the principals in this system.

The distinction between "the person in closest supervisory contact with the beginning teacher" and "superintendents and supervisors" as sources of information about beginning teachers had a number of procedural consequences for the study. First, this distinction revealed that substantial differences in the organization of supervisory services existed between systems, thus permitting elaboration of the organizational characteristics of the systems sampled. Systems C, I, K, L and M, placed the direct supervision of beginning teachers almost wholly with the principals. Systems, B, D, E, F, G, H and J divided direct supervision between the principals and a supervisor; within this group, however, some supervisors did much more direct supervision than others. Finally, in system A, a supervisory staff did virtually all of the direct supervision, while principals did very little, Table 4 presents a summary of the number of supervisors in each system, the ratio of supervisors to first year teachers during the two-year period covered by the study, and the approximate number of vis a vis contacts between supervisors and beginning teachers, organized according to the adjusted assessed valuation per resident pupil ADA.

Underlying Table 4 and relatively apparent in it is an association of direct relevance to the study. When systems were divided at the median adjusted assessed valuation, and the number of contacts per year between supervisor and teacher (4 or more contacts versus 3 or less contacts) were observed for primary and intermediate teachers separately, close associations between the wealth of the system and the number of contacts between supervisors and teachers (i.e. amount of supervision beyond that given by the principal) were found. In systems above the median in assessed valuation 34 of 37 intermediate teachers had four or more contacts per year with a supervisor, while in systems below the median in assessed valuation, only 15 of 32 intermediate teachers had more than 4 contacts. The resulting Chi Square value was 16.873, significant at $p < .001$. The relationships for primary teachers were similar, but considerably weaker, resulting in a Chi Square value of 5.610 significant at $p < .02$. These data firmly established an association between amount of supervision beyond that given by the principal and the adjusted assessed valuation of the school system. As will appear on later pages, this association is a critical one in interpreting the outcomes of the study.

TABLE 4. PROVISION OF SUPERVISOR SERVICES IN PARTICIPATING SYSTEMS ACCORDING TO SYSTEM WEALTH.

Systems ranked by Wealth	Number of Supervisors*	Ratio of Supervisors Per First Year Teacher	Approximate Number of Contacts With Each Beginning Teacher in First Year
G	1	1:9	4
A	7	1:2	18
J	1	1:4.5	9
K	0	0:4	0
E	2	1:6	4
D	1	1:13	2
Median			
B	1	1:13	6
F	1	1:10	18
C	1	1:13	0
L	0	0:4.5	0
I	0	0:4	0
H	1	1:3.5	3
M	0	0:1.5	0

*Superintendents not included.

Quantification of the amount of supervision given a beginning teacher by the person who had closest supervisory contact (typically the principal) was obtained by converting the number of quarter hours of supervision per month by this person for each teacher into a nine-point scale. Thus a teacher receiving 0-14 minutes supervision per month received a scale value of 1, while a teacher receiving 120 or more minutes of supervision per month received a scale value of 9. Because information relative to amount of supervision per month was obtained at the end of the 1962-63 school year, this information was not available for the first year of experience of teachers in the 1961 sample A summary of the mean and range of supervision received by teachers in each system from the person having closest supervisory contact, separated for the 1961 and 1962 samples and organized by the wealth of the system, appears in Table 5.

TABLE 5. MEAN AND RANGE OF NUMBER OF QUARTER HOURS OF SUPERVISION GIVEN THE 1961 AND 1962 SAMPLES, BY WEALTH OF SYSTEM.

Systems Ranked by Wealth	Supervision Received			
	1961 Sample		1962 Sample	
	$\bar{X}$	range	$\bar{X}$	range
G	6.00	3-9	4.79	1-9
A	4.30	2-9	4.93	2-9
J	3.33	2-5	4.40	2-9
K	3.00	2-5	6.25	3-9
E	3.62	1-9	4.67	2-7
D	3.75	2-7	3.00	1-7
Subsample $\bar{X}$	3.93		4.69	
B	3.00	2-4	3.11	1-9
F	5.50	5-6	6.00	2-9
C	6.30	2-9	6.63	2-9
L	3.33	3-4	4.25	2-7
I	3.25	2-5	3.75	3-5
H	3.50	2-6	3.00	2-4
M	--	--	--	--
Subsample $\bar{X}$	4.73		4.84	

The data shown in Table 5 are important in the study in two respects. First, they indicate that second year teachers received only slightly less supervision per month than first year teachers. On the average, first year teachers received about an hour and ten minutes supervision per month, while second year teachers received about an hour and five minutes. Because of these small differences, as well as to maintain sufficient cell frequencies in subsequent analyses, the 1961 and 1962 samples were pooled when amount of supervision was an independent variable in the statistical analyses. Second, there is little indication in Table 5 that the assessed valuation of the systems sampled was associated with the amount of supervision given by the person having closest supervisory contact. This fact made it possible to extract the effects of amount of supervision on teacher performance independent of the effects of the assessed valuation of the system in which the teacher taught.

The distinction between "the person having closest supervisory contact with the beginning teacher" and "superintendents and supervisors" not only affected the way in which supervision as a variable was dealt with, as described above, it also directly affected the operational definitions of "teaching success" as a criterion variable. Those persons who had closest supervisory contact with beginning teachers were asked to appraise teachers in each of five teaching areas on a three point scale, as previously described, at the end of both 1961-62 and 1962-63. This procedure led to a criterion of teaching success for all participating teachers except those who quit before completing a year of teaching. A second criterion of teaching success was the over-all appraisal of each beginning teacher still in the system by the system supervisor or superintendent at the close of the study in 1962-63. These procedures were to have led to two appraisals for most teachers; however, by the end of the study, it was clear that each variable had been compromised by the combined effect of design errors, which failed to take sample erosion adequately into account, leniency errors by raters and differences in the organization of supervisory services among school systems.

The design errors occurred through the investigator's failure to obtain an interview appraisal of teacher success by supervisors or superintendents of those teachers in the 1961 sample who quit during or at the end of the 1961-62 school year. Rather, this appraisal was obtained from the person who had closest supervisory contact with teachers. This error resulted in no supervisory or superintendent appraisal of success for the 25% of the 1961 sample who quit at the end of the first year.

The effect of this design error was greatly compounded by the observation that the appraisals of teaching success by the person having closest supervisory contact contained very great leniency errors. These errors varied by grade range and by school system. In most systems, teachers of grades 1-3 were rated higher than teachers of grades 4-6. Within grade levels, systems with low assessed valuation were more lenient than were those with high assessed valuation. The effects of these leniency errors, by grade level and by assessed valuation, may be observed in Table 6. The means shown in this table were obtained by assigning a weight of 4 to the high category, 2.5 to the middle category and 1 to the low category in each of the five areas of teaching appraised by the rater in his terminal rating of the teacher. The minimum possible rating was 5, the maximum, 20.

TABLE 6. VARIATION IN MEAN* RATINGS BY GRADE LEVEL, AND BY ASSESSED VALUATION OF SYSTEMS.

	Assessed Valuation#	
Grade Level	Above Median	Below Median
1-3	n=51 $\bar{X}=16.06$	n=58 $\bar{X}=17.23$
4-6	n=37 $\bar{X}=13.72$	n=29 $\bar{X}=15.34$

#The standard deviation within these groups was approximately 2.0.

From a procedural viewpoint, the leniency errors shown in Table 6 precluded the use of pooled ratings across grade levels and systems since teachers of grades 1-3 in low assessed systems dominated the upper end of the scale and teachers of grades 4-6 in high assessed systems the lower end. While equalization through a Z-score transformation of the ratings of each grade level-assessed valuation group was technically possible, this strategy proved to be unsuccessful because of variations between systems within these categories, which confounded leniency errors attributable to individual systems with the appraisal of performance. From the viewpoint of sample characteristics, the data in Table 6 are of interest in that it suggests that principals in low assessed valuation systems perceived their beginning teachers as being relatively good, while principals in high

---

[*] Throughout the paper $\bar{X}$ is used for the mean.

assessed valuation systems tended to perceive their teachers as being more nearly average or toward the middle of the distribution of teaching performance. While a number of hypotheses may be entertained to explain these differences, it is implicit in the investigator's interpretation of them as leniency errors, that raters in high assessed systems, especially when rating intermediate teachers, operated on a more stringent criterion concerning what constitutes successful performance than did raters in low assessed valuation systems.

In order to avoid the leniency errors associated with ratings by the person having closest supervisory contact, this criterion was largely, but not wholly, abandoned as a variable in the study. The appraisal of teacher performance by supervisors and superintendents was used largely in its place. The former criterion was not wholly abandoned because it represented the only estimate of teaching success for teachers who left their positions at the end of 1961 and for teachers in system C, in which the supervisor delegated all responsibilities for appraisal to the principals. In order to avoid losing a substantial number of subjects who quit, as well as those in one entire school system, in analyses involving estimates of teacher success, the principal's rating for these groups were merged with those made by supervisors and superintendents at the end of 1962-63 by the following procedure: appraisals by supervisors or superintendents were dichotomized for each participating system. Teachers appraised as average or below average were placed in one group, those appraised as above average or superior in the other. In order to place those teachers who quit at the end of 1961-62 into one or the other of these categories, the principal's ratings for all teachers within each system were dichotomized at the median for the system. The position in the distribution of each teacher who quit was then noted, and assignments made for those below the median for the system to the average-below average category, and for those above the median to the above average-superior category. The same procedure was utilized in assigning the teachers in system C to the average-below average and the above average-superior groups.

The average-below average group and the above average-superior groups in each system were subsequently collected into four cells: intermediate teachers in high assessed systems, primary teachers in high assessed system, intermediate teachers in low assessed systems and primary teachers in low assessed systems. Observation of the frequencies in the average-below average group and the above-average-superior group in each of these cells indicated that the procedures utilized had effectively removed leniency errors between grade levels in high assessed systems, yielding approximately equal numbers of teachers appraised as average-below average and above-average-superior in each grade range. In low assessed systems, however, the leniency errors persisted, with 3 of every 5 intermediate teachers and 2 of every 3 primary teachers being appraised as above average-superior.

This persistent tendency of systems with relatively low assessed valuation to perceive their teachers as superior had two consequences for the study. Procedurally, it permanently precluded the possibility of pooling teachers from high and low assessed valuation systems in analyses involving the appraisal of success. The preclusion of this possibility, plus the necessity to control for grade level effects, greatly reduced the cell frequencies in the factorial designs employed, thus weakening the statistical aspects of the study. Second, the tendency of supervisors and principals in low assessed systems to perceive their teachers as relatively superior suggests that they may have given a somewhat different kind of supervisory treatment to their teachers than was the case in the high assessed systems, thus making it advisable, if not necessary, to keep leniency differences between systems in mind in interpreting the results statistically attributable to the interaction between the assessed valuation of the system and the amount of supervision received by teachers.

In addition to relationships to leniency errors, and to supervision above and beyond that given by principals, the assessed valuation of the systems sampled was also linked to the proportions of working class and middle class children in the schools in the systems. The predominant class background of the children in each school building in each system was obtained from supervisors or superintendents by interview as previously indicated. The original scale on which these persons rated schools had three categories: lower class, lower and middle middle class, and upper middle class. During the interviews, however, supervisors and superintendents elaborated considerably on the class background of each school, especially distinguishing lower middle and middle middle class. On the basis of these distinctions, the three categories initially employed were subsequently collapsed to two, working class and middle class. Lower lower, upper lower, and lower middle class schools were included in the category "working class,", while middle middle, upper middle and upper class schools were included in the category "middle class".

Since school buildings vary considerably in the number of classroom units they contain, estimates of the proportions of working class to middle class children on a school by school basis is open to error attributable to this variation. In order to remove this source of error, the number of classroom units in each school in each system was obtained, and these units assigned to the category "working class", or "middle class" according to the category designation of the schools containing these units. A summary of the number of classroom units designated as "working class" and "middle class" for each system, and the number of beginning teachers teaching in working class and middle class schools, organized by the assessed valuation of the system is shown in Table 7.

TABLE 7. SOCIO-ECONOMIC CLASS (SEC) BACKGROUND OF CHILDREN IN CLASSROOM UNITS, BY WEALTH OF SYSTEM.

Systems Ranked by Wealth	Number of Classrooms by SEC Background		SEC Background of Rooms Taught by Beginning Teachers	
	Working	Middle	Working	Middle
G	58	55	7	11
A	176	180	22	6
J	17	46	0	8
K	24	30	3	5
E	78	79	11	6
D	82	94	5	9
Subtotal	435	484	48	45
B	24	173	3	18
F	0	92	0	20
C	101	57	13	14
L	20	26	2	5
I	27	37	7	1
H	44	59	2	6
M	12	25	0	1
Subtotal	288	469	27	65

The data shown in the two left-hand columns of Table 6 strongly suggest that a relatively high ratio of working class to middle class children is a characteristic of the system with high assessed valuation, while a high ratio (about 2 to 1) of middle class to working class children is characteristic of the low assessed systems. These ratios are not, however, perfectly consistent within each group of systems. This variation, especially as characterized by the deviation of system J in the high assessed group and system C in the low assessed group, suggest that in a broader sample of systems, two additional cells would appear, a cell comprised of high assessed systems with a high ratio of middle class to working class children, and a cell comprised of low assessed valuation systems with a high ratio of working class to middle class children. In the present sample, the relatively high ratio of working to middle class children in high assessed system appears to occur as follows. High assessed valuation per pupil is predicated largely on the degree of industrialization in the community. High industrialization is in turn associated with a relatively large working class employed in these industries. Low assessed valuation per pupil, on the other hand, is predicated on having predominantly residential property, which provides, at least in the present sample, a relatively limited tax base, but a relatively large number of middle class children. The deviation of system J and C from this general pattern appears to occur in an orderly way . System J is predominantly residential, and rural, but has within it a single wealth-producing plant, which, however, employs relatively few persons in the community, thus creating high assessed valuation without a concomitant working class. System C, on the other hand, is adjacent to a highly industrialized area, but is not in itself highly industrialized. Its residential property, however, is occupied in substantial part by the working class which serves the adjacent industrial area.* Thus, this system is saddled with a great many working class children, but a relatively limited tax base.

Observation that the socio-economic status of the children taught by beginning teachers varied with the assessed valuation of the systems in which they taught had a major consequence for the study. It made very clear that, as a variable, assessed valuation was broadly confounded with other variables which could not, in the present study, be fully extracted from it. That is, the amount of supervision received from supervisors, the leniency of principals and supervisors in appraising teachers, and the relative proportions of working class and middle class children all varied with assessed valuation.

On the basis of these observations the inference was made that teachers sampled were teaching in one or the other of two broad types of institutional context. One context was defined primarily by systems with relatively high assessed valuation per resident pupil

---

*Approximately one of every two workers residing in this community work outside the community. U.S.: Census of Population; 1960. Final report IRC (1)-16C, Indiana.

ADA, and characterized by: (1) the provision of substantial supervisory services above and beyond that provided by principals; (2) more stringent supervisory perception of the performance of beginning teachers and, (3) a relatively high proportion of working class to middle class children in the system. This type of institutional context was labeled "pattern A context". This pattern was principally associated with highly industrialized communities, and appears in most distinct form in such communities with population bases in excess of 30,000 persons.

The second context was defined primarily by systems with relatively low assessed valuation, and characterized by: (1)placement of the responsibilities for supervision primarily in the hands of the principals; (2) less stringent perception of the performance of beginning teachers; (3) a relatively high proportion of middle class to working class children in the system. This type of institutional context was labeled "pattern B context". Generally, this type of institutional context was identified with predominantly residential communities with a relatively low level of industrialization within the community.

While the systems included in the pattern A context as opposed to the pattern B clearly differed along some dimensions, along others they were quite similar. For example, the mean per pupil expenditures in 1960-61, as obtained from Indiana State Department of Education records was $325.25 for the systems with pattern A, and $337.76 for the systems in pattern B. The mean beginning teachers salary in 1960-61 in pattern A systems was $4,600 and in pattern B, $4,650, while the mean expected salary at the end of five years experience with the Masters Degree was $5,562 for pattern A systems, and $5,582 for pattern B systems. The variations in salaries occurred primarily within patterns rather than between patterns.

Beginning teachers in pattern A and pattern B systems averaged approximately 30 pupils per class, with large variation within systems and slight variation between systems. The hours per week each beginning teacher devoted to teaching arithmetic and to teaching reading in grades 4-6 was also similar between patterns A and B, with the average intermediate teacher devoting about 4 hours to teaching arithmetic and about 5 hours to teaching reading. The amount of formal in-service activity directly devoted to teaching reading in pattern A and pattern B systems was very similar, with only three teachers in each pattern involved in this sort of activity during the two year period covered by the study. In-service activity directly related to teaching arithmetic, however, was slightly different between systems with seven pattern B teachers, grades 4-6, participating in a workshop or college course on arithmetic, but only 2 teachers thus participating in pattern A systems. Generally speaking, participating systems appeared to feel that a high level of involvement in formal in-service work was undesirable for beginning teachers, especially during the first year of experience.

The two patterns of institutional contexts utilized in the study should be clearly recognized as "empirically" or "inductively" arrived at patterns. They were not identified on theoretical grounds before the study began; rather, they were identified through examining relationships among a number of variables, which, at the outset, were defined only as variables relevant to the institutional context of teaching. In spite of this somewhat after the fact character of the patterns, their utilization was critical in the grouping of data for statistical analysis and for the interpretation of results.

CHAPTER III

PROCEDURAL DETAILS AND RESULTS

Variables Influencing Performance Change

The variables anticipated to effect performance change on the MTT and the TTR fall into three groups: (1) those intrinsic to the institutional context in which the teacher taught, including grade level, system type, amount of supervision received, in-service training and hours per week class schedules in reading and arithmetic; (2) years of teaching experience; and (3) variables intrinsic to teachers, including beginning scores on the MTT and TTR, and scores on the TCS scales. Among these variables grade level taught and system type had a pervasive influence in the sense that control of them was essential in extracting other results. Hence, these variables are dealt with first.

Effects of grade level and system type. To determine these effects, teachers in the 1961 and 1962 samples were pooled, with all those who left their positions prior to the post-test excluded. The remaining teachers were divided according to whether they taught in pattern A or pattern B systems, and redivided by grade ranges 1-3 and 4-6. These groups were then used in two,three-way repeated measures analyses of variance, with MTT scores functioning as the dependent variable in one analysis and TTR scores as the dependent variable in the other.

In Table 8, the raw score means for pre-test and post-test scores by grade level and system type for the MTT are arrayed, followed in Table 9 by the analysis of variance. In Table 10, the raw score means for pre-test and post-test scores by grade level and system type for the TTR are arrayed, followed in Table 11 by the analysis of variance. Because the analysis of variance used is extremely sensitive to disproportionality, three subjects were randomly deleted from the grades 1-3 type B system cell in the analysis of MTT scores, and six subjects from this cell were randomly deleted in the analysis of the TTR scores. These deletions brought the cell frequencies into exact proportionality.

In the first part of these analyses, the main effects for grade level reflect differences in the pooled pre-test, post-test means by grade level; the main effects for system type reflect differences in the pooled pre-test, post-test means by system type, while the first order interaction, system type of grade level, reflects the differential differences between pooled pre-test post-test means by grade and system type. The significance of these effects are estimated against an error term (error 1) which reflects the random variation within each pooled pre-test, post-test cell; this error term is comparable to that obtained in a standard two-way analysis of variance. In the second part of these analyses, the main effect for test (gain) reflect the differences in pre-test and post-test scores, grade level and system type pooled; the first order interaction, system type by test, reflects differences between system types by differences between pre-test and post-test scores, grade levels pooled; the first order interaction grade by test, reflects differences between grades by differences between pre-test and post-test scores, system types pooled, while the second order interaction, grade by system type by test, reflects differences in the difference scores between pre-test and post-test, by grade level and system type. The significance of these effects are estimated against an error term (error 2) which reflects random variation of the difference score for each subject (teacher) around the mean difference score between pre-test and post-test for each grade level-system type cell.

TABLE 8. MEANS FOR PRE-TEST AND POST-TEST MTT SCORES BY SYSTEM TYPE AND GRADE LEVEL

	Pattern A Systems		Pattern B Systems	
	Primary	Intermediate	Primary	Intermediate
	n = 46	n = 31	n = 36	n = 24
Pre-test mean	21.09	23.65	21.00	21.16
Post-test mean	22.26	27.10	21.11	23.21

Examination of Tables 8 and 10, and the respective analyses shown in Tables 9 and 11 indicate clearly that there were significant increases in performance on both the MTT and the TTR between pre-test and post-tests. These increases, however, were not distributed in the same way on the MTT and TTR over the groups involved. The increases on the TTR were equally distributed over the groups, thus producing significant main effects for the tests, but no significant interactions. The increases on the MTT, on the other hand, were not equally distributed over the groups, but were concentrated in grades 4-6 cells, which appear to have produced both the significant main effect for tests and a significant grade x test interaction. In addition, Table 9 shows that the pooled pre-test, post-test scores on the MTT differ significantly by grade level and by system. These differences were concentrated in the relatively high pre-test and post-test means of intermediate teachers in pattern A systems.

TABLE 9. ANALYSIS OF VARIANCE FOR PRE-TEST AND POST-TEST MTT SCORES BY SYSTEM TYPE AND GRADE LEVEL

Source	SS	df	MS	F	p
Grade level	438.05	1	438.05	10.24	.005
System type	184.51	1	184.51	4.32	.05
System x grade	105.11	1	105.11	2.46	n.s.
Error 1	5686.78	133	42.76		
Tests	167.14	1	167.14	9.33	.005
Tests x system	24.56	1	24.56	1.37	n.s.
Tests x grade	74.61	1	74.61	4.38	.05
Tests x grade x system	.28	1	.28	<1	n.s.
Error 2	2382.41	133	17.91		

TABLE 10. MEANS FOR PRE-TEST AND POST-TEST TTR SCORES BY SYSTEM TYPE AND GRADE LEVEL.

	Pattern A Systems		Pattern B Systems	
	Primary	Intermediate	Primary	Intermediate
	n = 44	n = 31	n = 34	n = 24
Pre-test mean	17.91	18.23	18.50	16.75
Post-test mean	21.27	22.29	21.29	21.04

TABLE 11, ANALYSIS OF VARIANCE FOR PRE-TEST AND POST-TEST TTR SCORES BY SYSTEM TYPE AND GRADE LEVEL

Source	SS	df	MS	F	p
Grade level	0.29	1	0.29	<1	n.s.
System type	9.69	1	9.69	<1	n.s.
Grade X system	44.04	1	44.04	<1	n.s.
Error 1	5751.98	129	44.58		
Tests	837.59	1	837.59	51.45	.001
Tests x system	0.86	1	0.86	<1	n.s.
Tests x grade	17.66	1	17.66	1.08	n.s.
Tests x grade x system	20.33	1	20.33	1.25	n.s.
Error 2	2100.56	129	16.28		

These findings suggest that performance on the MTT was subject to variation by grade level and by system type. Thus, separation of grade level and system groups was maintained in analyses involving other independent variables. The TTR was not subject to the same sources of variation, but separation of grade level and system was maintained for this variable, nonetheless, primarily on intuitive grounds.

Effects of Supervision. "Supervision" was defined as the amount of supervision given by that person in each system having closest supervisory contact with each beginning teacher. As indicated in the preceding chapter, this variable was scaled from 1-9 on the basis of the quarter hours of supervision given per month. For the analyses which follow, the variable was dichotomized as one hour or less per month (low supervision) versus more than one hour per month (high supervision).

Because the number of independent and dependent variables to be considered in these analyses make it difficult to see the main pattern in the abundance of detail, the method of organization will be explained first. Differences in the effects of supervision on the MTT scores of intermediate teachers according to system type are shown first (Tables 12-15), immediately followed by the same comparisons but with the TTR as the dependent variable (Tables 16-19). Next, the same comparisons are made using primary rather than intermediate teachers, with the analysis utilizing the MTT as the dependent variable coming first (Tables 20-23), and the analysis utilizing the TTR as the dependent variable, second (Tables 24-27). This organization emphasized differences between system types for teachers in the same grade ranges, rather than differences between grade ranges within system types.

In Table 12 are shown the mean MTT raw scores on pre-test and post-test for teachers of grades 4-6 in pattern A systems. Table 13 shows the two-way repeated measures analysis of variance for this table. In Table 14 are shown the mean MTT raw scores on pre-test and post-test for teachers, grades 4-6, in pattern B systems, while Table 15 shows the analysis of variance of these means.

TABLE 12. MEANS FOR PRE-TEST AND POST-TEST MTT SCORES FOR INTERMEDIATE TEACHERS IN PATTERN A SYSTEMS BY AMOUNT OF SUPERVISION RECEIVED.

	High Supervision		Low Supervision	
	Pre-test	**Post-test**	Pre-test	Post-test
n	13	13	18	18
$\bar{X}$	21.69	28.00	25.06	26.44

TABLE 13. ANALYSIS OF VARIANCE OF MEANS IN TABLE 12.

Source	SS	df	MS	F	p
Supervision	12.49	1	12.49	< 1	n.s.
Error 1	1163.63	29	40.13		
Tests	184.82	1	184.82	10.44	.005
Supervision x tests	91.15	1	91.15	5.15	.05
Error 2	513.53	29	17.71		

TABLE 14. MEANS FOR PRE-TEST AND POST-TEST MTT SCORES FOR INTERMEDIATE TEACHERS IN PATTERN B SYSTEMS BY AMOUNT OF SUPERVISION RECEIVED.

	High Supervision		Low Supervision	
	Pre-test	Post-test	Pre-test	Post-test
n	11	11	13	13
$\bar{X}$	20.27	22.09	21.92	24.15

TABLE 15. ANALYSIS OF VARIANCE OF MEANS IN TABLE 14.

Source	SS	df	MS	F	p
Supervision	41.08	1	41.03	<1	n.s.
Error 1	937.74	22	46.62		
Tests	50.03	1	50.03	3.11	n.s.
Tests x supervision	.51	1	.51	<1	n.s.
Error 2	353.96	22	16.09		

As shown in Tables 13 and 15, the effect of supervision on increases in performance on the MTT depends on the type of system in which the teacher taught. As is apparent in Tables 12 and 13, increases in MTT performance are concentrated among intermediate teachers who taught in pattern A systems and who received more than one hour per month supervision. The effect in this group is very strong, and serves not only to produce a significant test x supervision interaction, but also large main effects for tests. These effects are wholly absent among the intermediate teachers in pattern B systems, on whom supervision had no measurable effect and who did not significantly increase in MTT scores between testings.

A similar, but not identical, effect appeared when the TTR rather than the MTT was utilized as the dependent variable. These effects may be examined in Tables 16 through 19.

TABLE 16. MEANS FOR PRE-TEST AND POST-TEST TTR SCORES FOR INTERMEDIATE TEACHERS IN PATTERN A SYSTEMS BY AMOUNT OF SUPERVISION RECEIVED.

	High Supervision		Low Supervision	
	Pre-test	Post-test	Pre-test	Post-test
n	14	14	17	17
$\overline{X}$	16.57	21.93	19.59	22.59

TABLE 17. ANALYSIS OF VARIANCE OF MEANS IN TABLE 16.

Source	SS	df	MS	F	p
Supervision	51.89	1	51.89	1.10	n.s.
Error 1	1371.99	29	47.31		
Tests	256.07	1	256.07	48.04	.001
Tests x supervision	91.32	1	91.32	17.13	.001
Error 2	154.61	29	5.33		

TABLE 18. MEANS FOR PRE-TEST AND POST-TEST TTR SCORES FOR INTERMEDIATE TEACHERS IN PATTERN B SYSTEMS BY AMOUNT OF SUPERVISION RECEIVED.

	High Supervision		Low Supervision	
	Pre-test	Post-test	Pre-test	Post-test
n	11	11	13	13
$\overline{X}$	15.00	19.00	18.23	22.77

TABLE 19. ANALYSIS OF VARIANCE FOR MEANS IN TABLE 18.

Source	SS	df	MS	F	p
Supervision	145.98	1	145.98	2.92	n.s.
Error 1	1098.00	22	89.91		
Tests	221.02	1	221.02	56.09	.001
Tests x supervision	.86	1	.86	<1	n.s.
Error 2	86.82	22	3.92		

In pattern A systems, the effects of supervision on TTR performance are similar to those on the MTT. There is, however, one difference; both high and low supervision groups gain, but with the greater gains coming in the high supervision group. It thus appears to be the case that the effects for supervision are somewhat independent of the main effects for tests when TTR is the dependent variable, while when MTT scores are the dependent variable, it is the interaction between supervision and tests which carry the main effect gains for tests. These results imply that some variable in addition to supervision affects gains on the TTR, while it is primarily supervision which affects gains on the MTT in the pattern A systems.

In pattern B systems, (Table 19) the presence of a large main effect for tests but the absence of an interaction between tests and supervision also indicates the presence of at least one variable which leads to gains, independent of the amount of supervision received. The fact that this effect is about equally present in both pattern A and pattern B systems suggests the possibility that the variable involved is classroom experience in teaching. This possibility is further examined at a later point.

Changes in the performance of primary teachers on the MTT was of particular interest in the study since the MTT was considered irrelevant to the type of arithmetic taught in the primary grades. This group therefore functioned as a control group for intermediate teachers, although not a perfect control, since primary teachers do teach some arithmetic. The changes in MTT scores for this group, by amount of supervision, may be examined in Table 20 through 23.

TABLE 20. MEANS FOR PRE-TEST AND POST-TEST MTT SCORES FOR PRIMARY TEACHERS IN PATTERN A SYSTEMS BY AMOUNT OF SUPERVISION RECEIVED.

	High Supervision		Low Supervision	
	Pre-test	Post-test	Pre-test	Post-test
n	17	17	29	29
$\overline{X}$	23.12	23.71	19.90	21.41

TABLE 21. ANALYSIS OF VARIANCE FOR MEANS IN TABLE 20.

Source	SS	df	MS	F	p
Supervision	162.87	1	162.87	4.87	.05
Error 1	1559.35	44	35.44		
Tests	31.70	1	31.70	1.83	n.s.
Tests x supervision	4.63	1	4.63	<1	n.s.
Error 2	763.67	44	17.36		

TABLE 22. MEANS FOR PRE-TEST AND POST-TEST MTT SCORES FOR PRIMARY TEACHERS IN PATTERN B SYSTEMS BY AMOUNT OF SUPERVISION RECEIVED.

	High Supervision		Low Supervision	
	Pre-test	Post-test	Pre-test	Post-test
n	21	21	18	18
$\overline{X}$	21.43	20.52	20.33	21.06

TABLE 23. ANALYSIS OF VARIANCE FOR MEANS IN TABLE 22.

Source	SS	df	MS	F	p
Supervision	1.54	1	1.54	<1	n.s.
Error 1	1947.69	37	52.64		
Tests	.47	1	.47	<1	n.s.
Tests x supervision	12.77	1	12.77	<1	n.s.
Error 2	609.69	37	16.48		

As indicated in Tables 21 and 23, primary teachers do not increase their scores on the MTT with teaching experience, irrespective of amount of supervision, and irrespective of system type. The skills measured by the MTT are relevant only to intermediate teachers. Moreover, the fact that primary teachers did not significantly improve their pre-test scores strongly suggests that taking the MTT twice with long intervals between tests does not result in significant practice effects.

The appearance in Table 21 of a significant effect for supervision (pooled pre-test, post-test) was an unexpected result. This effect may be an alpha error or it may in fact be that a high MTT score among primary teachers in pattern A systems leads to a higher degree of supervision. This finding is appraised further in the immediately succeeding section in which those variables which lead to greater or less supervision are examined.

Changes in the performance of primary teachers on the TTR were quite different from those on the MTT. The nature of these changes and their statistical significance by type of school system may be observed in Tables 24 through 27.

The significance of the main effects for tests (gains) as shown in Tables 25 and 27, but the non-significance of the test by supervision interactions indicate that primary teachers increased in score between tests, but that this increase is not specifically attributable to the amount of supervision received. Rather, it appears to be attributable to classroom experience in teaching reading as was suggested for the main effect gains in TTR scores among intermediate teachers.

TABLE 24. MEANS FOR PRE-TEST AND POST-TEST TTR SCORES FOR PRIMARY TEACHERS IN PATTERN A SYSTEMS BY AMOUNT OF SUPERVISION RECEIVED.

	High Supervision		Low Supervision	
	Pre-test	Post-test	Pre-test	Post-test
n	16	16	28	28
$\overline{X}$	17.56	20.31	18.12	21.82

TABLE 25. ANALYSIS OF VARIANCE FOR MEANS IN TABLE 24.

Source	SS	df	MS	F	p
Supervision	21.47	1	21.47	<1	n.s.
Error 1	1785.31	42	42.51	12.36	.005
Tests	248.91	1	248.91		
Tests x supervision	4.73	1	4.73	<1	n.s.
Error 2	845.86	42	20.14		

TABLE 26. MEANS FOR PRE-TEST AND POST-TEST TTR SCORES FOR PRIMARY TEACHERS IN PATTERN B SYSTEMS BY AMOUNT OF SUPERVISION RECEIVED.

	High Supervision		Low Supervision	
	Pre-test	Post-test	Pre-test	Post-test
n	21	21	19	19
$\overline{X}$	17.52	21.91	19.26	21.05

TABLE 27. ANALYSIS OF VARIANCE OF MEANS IN TABLE 27.

Source	SS	df	MS	F	p
Supervision	3.92	1	3.92	<1	n.s.
Error 1	1623.63	38	42.73		
Tests	198.45	1	198.45	9.64	.01
Tests x supervision	33.50	1	33.50	1.63	n.s.
Error 2	782.05	38	20.58		

In order to attribute the main effect gains on the TTR over all of the intermediate and primary teachers as shown in Tables 17, 19, 25 and 27 to experience in teaching reading, a control group who taught grades 1-6, but did not teach reading would be necessary. A control group meeting these exact conditions was not available; however, eight teachers who taught kindergarten, art or physical education and who had no experience teaching reading in grades 1-6 were available from the group excluded from the main body of the study. While this group is not fully adequate as a control group in terms of size, the pre-test and post-test TTR scores for the group do serve as a relevent reference point. The pre-test mean was 15.25, and the post-test mean, 15.12; thus, this group made no gain in the TTR Score. The fact that it did not gain suggests that the gains on the TTR among teachers who taught reading in grades 1-6 could reasonably be attributed to the effects of experience in teaching reading.

The attribution of score increase on the TTR to experience in teaching reading marks a fundamental difference between the variables affecting increases on the TTR and those affecting increases on the MTT. Only amount of supervision received was shown to be associated with increases in MTT scores, and then only among intermediate teachers in pattern A systems, while experience in teaching reading appears to be the principal variable influencing increases in TTR scores, with supervision providing considerable added effect among intermediate teachers in pattern A systems.

Effects of In-Service Work. Direct participation in formal in-service activity (workshops and college courses) related to teaching either reading or arithmetic was obtained from the statements of beginning teachers on the day the post-tests were given. Of the total group of teachers involved in the study, only six, three from pattern A systems and three from pattern B systems had formal work in language arts or reading, while nine teachers did work relevant to teaching arithmetic, two from type A systems and seven from type B system, all in grades 4-6. Because of the small number of teachers involved in formal in-service activity, only one analysis was made. In it, the seven teachers in pattern B system who did in-service work relevant to teaching arithmetic were compared to the 17 who did not do such work. Tables 28 and 29 give the data relevant to this comparison.

TABLE 28. MEANS OF PRE-TEST AND POST-TEST MTT SCORES FOR INTERMEDIATE TEACHERS PARTICIPATING AND NOT PARTICIPATING IN IN-SERVICE ACTIVITY RELATED TO TEACHING ARITHMETIC IN PATTERN B SYSTEMS.

	Participating		Not Participating	
	Pre-test	Post-test	Pre-test	Post-test
n	7	7	17	17
$\bar{X}$	19.29	24.57	21.94	22.67

TABLE 29. ANALYSIS OF VARIANCE OF MEANS IN TABLE 28.

Source	SS	df	MS	F	p
Participation	1.33	1	1.33	<1	n.s.
Error 1	977.49	22	44.43		
Tests	50.03	1	50.03	5.06	.05
Tests x participation	137.35	1	137.35	13.92	.005
Error 2	217.12	22	9.87		

Tables 28 and 29 are revealing on two counts. First, participation in relevant in-service activity by intermediate teachers in pattern B systems does have a significant effect on scores increases on the MTT. Second, comparison of Table 29 to Table 14, for which the sum of squares are identical, indicates that the main effects for tests is significant in Table 29, but not in Table 14. The reason for this discrepancy is that the interaction between test and participation is shown in Table 29 absorbs sufficient variance to greatly reduce error 2, thus permitting the emergence of significant main effects. These effects, however, as may be observed by examination of Table 28, come about primarily through the much greater score increase of the teachers participating in in-service work, which increase carries, as it were, the main effects. Thus, the proper interpretation of Table 29 is that of an interaction-dominated main effect, rather than that of independent main effect and interaction. It was participation in in-service activity related to teaching arithmetic, then, which accounted for a substantial proportion of the total score increases in MTT performance among intermediate teachers in pattern B systems.

Effects of hours per week class schedule. The hours per week each participant spent teaching arithmetic and reading respectively were obtained from the principals. The hours per week spent teaching arithmetic was a relevant variable in relation to changes in MTT scores primarily for intermediate teachers. This variable was closely distributed among this group, however, and when dichotomized produced relatively little difference between the resulting two groups in actual time per week spent teaching arithmetic. The group spending the lesser time spent 45-50 minutes per day (3½-4 hrs./wk.) while the group spending greater time spent 55-60 minutes per day (4½-5 hrs./wk.); subsequent analysis of variance of changes in MTT scores between the group spending the lesser and the group spending the greater time, with system type controlled, indicated that hours per week class schedule was not a factor in changes in performance.

Hours per week spent teaching reading was better distributed among teachers than was hours per week spent teaching arithmetic, but two disturbing factors appeared in the use of this variable. The first was that time spent teaching reading was correlated with grade level taught, primary teachers spending the greater time, and second, that the amount of supervision received by intermediate teachers in pattern A systems was inextricably confounded with hours spent teaching reading.

To control for grade level effects, intermediate and primary teachers were separated. Intermediate teachers were then separated by system type, and within system types, teachers were dichotomized according to whether they taught reading five or less hours per week or more than five hours per week. The results by system type may be observed in Tables 30 through 33.

TABLE 30. MEANS OF PRE-TEST AND POST-TEST TTR SCORES FOR INTERMEDIATE TEACHERS BY HOURS PER WEEK CLASS SCHEDULE IN TEACHING READING IN PATTERN A SYSTEMS

	Five or Less Hours		More Than Five Hours	
	Pre-test	Post-test	Pre-test	Post-test
n	15	15	16	16
$\overline{X}$	19.87	21.80	16.69	22.75

TABLE 31. ANALYSIS OF VARIANCE OF MEANS IN TABLE 30.

Sources	SS	df	MS	F	p
Hours per week	19.24	1	19.24	<1	n.s.
Error 1	1404.64	29	48.44		
Tests	256.07	1	256.07	41.44	.005
Hours per wk. x tests	65.99	1	65.99	10.68	.005
Error 2	179.94	29	6.18		

TABLE 32. MEANS OF PRE-TEST AND POST-TEST TTR SCORES FOR INTERMEDIATE TEACHERS BY HOURS PER WEEK CLASS SCHEDULES IN TEACHING READING IN PATTERN B SYSTEMS.

	Five or Less Hours		More Than Five Hours	
	Pre-test	Post-test	Pre-test	Post-test
n	14	14	10	10
$\overline{X}$	17.43	21.71	15.80	20.10

TABLE 33. ANALYSIS OF VARIANCE OF MEANS IN TABLE 32.

Source	SS	df	MS	F	p
Hours per week	30.67	1	30.67	<1	n.s.
Error 1	1213.31	22	41.84		
Tests	221.02	1	221.02	55.53	.001
Hours per week x tests	.00	1	.00	<1	n.s.
Error 2	87.48	22	3.98		

According to Table 31, which contains a significant interaction, hours per week spent teaching reading influences gains on the TTR. Further examination of the characteristics of the teachers entering this analysis, however, indicated that of the 15 teachers who taught reading five or less hours per week, ten received low supervision, and five high supervision. Of the 16 teachers who taught reading more than five hours per week, on the other hand, nine received high supervision and seven low supervision. Hence, in this analysis, the effects of receiving high or low supervision are mixed with whatever effects hours per week class schedule in teaching reading might have. Observation of Table 16, which gives the effects of supervision on the same subjects as appear in Table 31, strongly suggests that what appears in the latter table is a weak supervision effect rather than a strong hours per week class schedule effect. Furthermore if hours per week class schedule were a strong effect, one might expect the effect to replicate in pattern B systems, which it did not, as shown in Table 33.

This state of affairs reflected a more general problem in making further analysis of increases in task performance, viz., that the number of variables requiring control had become too great for the number of subjects in the study, thus creating confounded effects. Further statistical analysis of score changes on the MTT and the TTR was therefore terminated at this point.

Patterns of Score Change by Months of Teaching Experience. In preceding sections, a number of variables were shown to influence changes in performance on the MTT and TTR. In this section all variables except grade level are disregarded in order to show the patterns of performance change across months or experience. No attempt was made to precisely appraise the patterns from a statistical viewpoint, for reasons pointed out in the preceding paragraph; rather, the intent here is to indicate the general changes that might be anticipated if one knew at the time a teacher graduated from his preparatory institution only what grade level he would be teaching during the next year or next two years. To facilitate graphic representation, the changes are shown in Z score form in Figure 1.

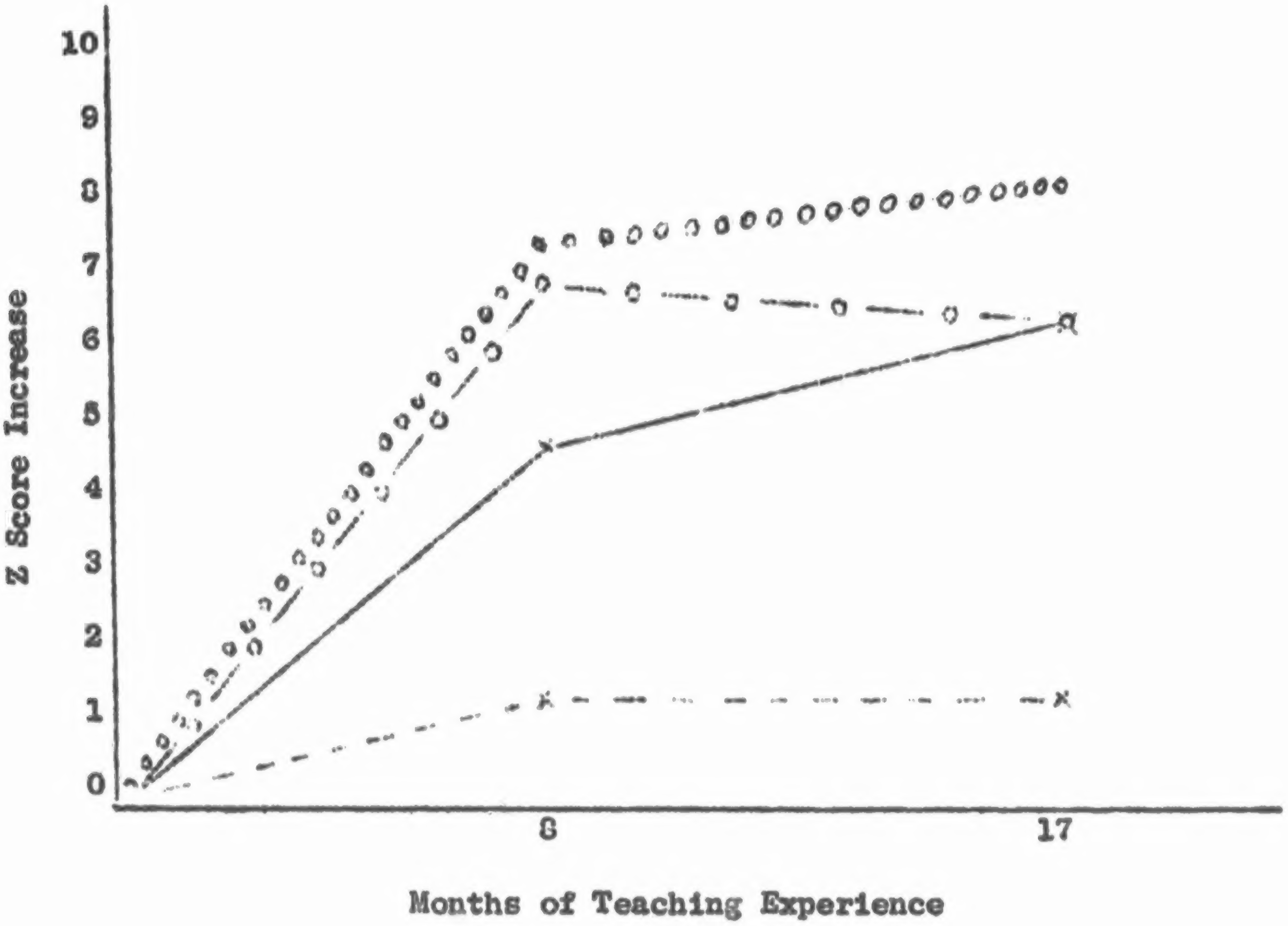

Figure 1. Changes in MTT and TTR Scores by Months of Teaching Experience.

Key
________ grades 4-6, MTT
- - - - - grades 1-3, MTT
0 0 0 0 0 grades 4-6, TTR
0-0-0-0-0 grades 1-3, TTR

Figure 1 distinctly suggests that most of the score increases on the MTT and the TTR come during the first year of experience, with a leveling off during the second year of experience. As a statistical point of reference, the square root of the "error 2" terms in Tables 9 and 11, translated into Z score form, give an approximate standard deviation against which to judge the degree of score change. In this case, standard deviation of about 6 is appropriate for comparing points on the same line. It might be noted that the small differences between the 1962 sample (8 months experience) and 1961 sample (17 months experience) underlay the pooling of these samples for purposes of statistical analysis in the study as a whole.

Effects of Pre-Test Levels and TCS Factors on Performance Change. The relationships among change scores on the MTT and TTR, pre-test levels on these instruments and TCS scores was examined via product moment correlation, with grade levels separate, but system types pooled. The results indicated that none of the TCS factors were directly related to performance change in either grades 1-3 or grades 4-6.* Pre-test scores on the MTT were related to MTT gains, however, ($r = -.66$, grades 4-6, $r = -.55$, grades 1-3), and pre-test scores on the TTR were related to TTR gains ($r = -.34$, grades 4-6, $r = -.55$, grades 1-3). These results indicated the presence of both regression effects and ceiling effects in the gain scores. While these effects did not have a direct influence on the results reported in previous sections, they probably did have an indirect effect on certain aspects of these results, as indicated below.

Analysis of the relationships between change scores on the MTT and the TTR, and amount of supervision received, with groups separated by grade level and system type, indicated that the amount of supervision received was negatively related to pre-test scores among intermediate teachers. The relationship of MTT pre-test to supervision among intermediate teachers in pattern A systems was -.32 ($p < .10$) and of TTR pre-test to supervision -.35 ($p < .10$). Among intermediate teachers in pattern B systems, the relationship of MTT pre-test to supervision was -.18 (n.s.) and TTR pre-test to supervision -.36 ($p = .10$). While these relationships are relatively weak, their presence is clearly suggested by examination of the pre-test means for high and low supervision groups in Tables 12, 14, 16 and 18. These effects, it might be noted, are most consistent in pattern A systems, in which amount of supervision is directly related to performance change, and least consistent in pattern B systems in which amount of supervision has no distinguishable effects on performance. Hence, low scores on the MTT and TTR pre-tests appear to be one factor associated with receiving a relatively high degree of supervision among intermediate teachers in pattern A systems.

---

* See Tables 36 and 37 on pages 68 and 69.

In addition to the level of pre-test scores on the MTT and TTR, two other variables were associated with the amount of supervision received by intermediate teachers in pattern A systems; one was factor B (traditional vs. permissive), $r = .40$ ($p < .05$) and the other, factor Z (stimulating, imaginative vs. dull, routine), $r = .37$ ($p < .06$). The higher scores on factor B indicate a more traditional viewpoint; thus the more traditionally oriented teacher received the greater supervision, while low scores on Factor Z are associated with unimaginative, routine teachers, who, in this case, were also more likely to receive the greater supervision. It thus appears that both low initial performance on the MTT and TTR, and certain personal-social characteristics have a bearing, through supervision, as a mediating variable, on changes in performance on the MTT and TTR among intermediate teachers in pattern A systems. It was this group, it might be noted, whose rather large performance changes were closely associated with the amount of supervision received.

TCS Variables As Remote Influences On Performance Change. As observed in the preceding section, factors B and Z from the TCS were associated with the amount of supervision received. The influence of these variables on performance change was thus once removed. Similarly, examination of the presence or absence of certain types of difficulties experienced by beginning teachers were shown to be associated with the amount of supervision received, and certain TCS variables were found to be associated with the presence of these difficulties, as described below.

The presence or absence of difficulties in teaching were determined from the responses of principals and supervisors to questions asking for a statement about the problems and difficulties experience by each teacher. These responses were diverse in content, however, three general categories were identified: (1) the absence of difficulties; (2) the presence of difficulties in discipline, classroom control, class management or classroom organization; and (3) the presence of difficulties in teaching a subject area, in public relations, or in general adjustment to teaching. The content of the problems in this latter category was diverse, but the presence of some sort of non-discipline related problem was clear. This category was labeled as a "miscellaneous difficulties" category.

Among these categories interesting differences between pattern A and pattern B systems appeared with respect to discipline, organization, management and control problems. Principals and supervisors in pattern A systems focused sharply on discipline and organization problems. Principals and superintendents in pattern B systems, on the other hand, did not mention discipline problems as such, nor did they stress organization, rather, most of the difficulties of teachers were expressed as "management" difficulties. Thus, within this category differences in the expression of teacher problems, if not also differences in their conceptualization, occurred between system types, a factor to be kept in mind in viewing subsequent results.

Analysis of the amount of supervision received by teachers in each of the above categories, with grade levels pooled but system types separated, indicated that teachers with discipline and/or classroom organization difficulties in pattern A systems received about one-half hour more supervision per month than either of the remaining two groups. This difference was significant at $p < .005$. Teachers with discipline and/or organizational problems in pattern B systems, however, did not receive significantly more supervision than the other two groups. The means underlying these findings may be observed in Table 34.

TABLE 34. QUARTER HOURS OF SUPERVISION RECEIVED BY GROUPS CATEGORIZED BY TYPES OF DIFFICULTIES IN TEACHING, BY SYSTEM TYPE.

	No Difficulties		Miscellaneous Difficulties		Discipline and Organizational Difficulties	
	n	$\bar{X}$	n	$\bar{X}$	n	$\bar{X}$
Pattern A systems	23	3.87	29	3.55	21	5.86
Pattern B systems	23	4.13	33	4.91	12	5.75

In the present context, these findings are significant primarily as an indication that it was the presence of discipline and/or organizational problems which constituted a major factor in the amount of supervision received by teachers in pattern A systems. This is not true however, in pattern B systems, in which the amount of supervision received was apparently a function of variables unlike those that lead to greater supervision in pattern A systems.

Subsequent to examination of the relationship between difficulties in teaching and amount of supervision received, the association between teaching difficulties and the various factors on the TCS were observed. The associations in this instance were predictive, since measurement of teachers on the TCS occurrred before the teachers experienced their difficulties. In pattern A systems, TCS factors X (warm, friendly), Y (responsible, business-like) and B (traditional vs. permissive) all distinguished at $p < .05$ between teachers who experienced no difficulties and teachers with discipline and/or organizational difficulties. In addition, factor B distinguished between those with no difficulties and those with miscellaneous difficulties. The presence of these differences suggested a fairly distinct pattern of personal-social variables for the person having difficulties with discipline and/or organization in systems of this type. These persons were relatively cool and aloof (low X), somewhat slipshod (low Y), but relatively traditional (high B). Of the variables contributing to this pattern, factor B seems most relevant since a relatively traditional viewpoint was associated not only with the presence of teaching difficulties, but also directly associated with a relatively high level of supervision. Factors X and Y, on the other hand, were specifically associated with discipline and organizational problems, but not directly associated with the presence of other types of difficulties or directly associated with the amount of supervision received.

The above findings were important to the study in two ways. First, they indicated that factors X and Y may have held remote or twice removed influences on performance change, since these two factors were linked to difficulties with discipline and organization which was in turn linked to supervision, which was directly associated with performance change. Second, the findings indicated that in pattern A systems, certain personal-social variables, viz., factors B, X, and Y predicted specific difficulties in teaching, viz., discipline and/or organizational problems, as seen by supervisory personnel.

Examination of the relationships between TCS factors and the presence or absence of teaching difficulties in pattern B systems yielded no statistically significant results. Insofar as the categories of teaching difficulties utilized were not related to the amount of supervision received in systems of this type, this finding was not surprising. It nonetheless underscores the fact that in the present study, no variables could be identified which would help explain why some teachers in pattern B systems receive more supervision than others. Statistically speaking, all the variance of the amount of supervision received by teachers in pattern B systems remains error variance.

Variables Associated With Supervisory Appraisals of Teacher Success

Variables associated with supervisory appraisals of success fall into three principal groups: those variables such as amount of supervision received and types of teaching difficulties encountered which may be thought of as constituting restricted criteria of success, and those independent variables, such as task performance and the TCS factors, which may be though of as contributing to success. In addition to these variables, system type, and within system type, the socio-economic class background of each school may be thought of as relevant control variables. These latter variables function as control variables in the sense that they moderate the associations between the independent variables and the criterion.

Relationships Among Criteria of Success. While in preceding sections neither the amount of supervision received nor the difficulties experienced by beginning teachers were utilized as criteria of success, each of these variables may in a limited way be thus utilized. Since the amount of supervision received by beginning teachers varied considerably from individual to individual, some teachers may be conceptualized not only to have received more supervision, but also to have "needed" more from the viewpoint of supervisory personnel. Those who "needed" more supervision might correspondingly be seen as being less successful in their first year to two of experience than those who needed less supervision. Thus an association between the amount of supervision received and supervisory appraisals of success might be anticipated. Examination of the amount of supervision received by teachers appraised as average-below average ("low" group) as opposed to those appraised as above average - superior ("high" group) by supervisors revealed, however, that no significant association existed between these two criteria. In both pattern A and pattern B systems the low group received only about 15 minutes more supervision per month than the high group. This difference was not significant at $p = .05$ by analysis of variance.

Observation of the types of difficulties experienced by beginning teachers in relation to appraisals of success indicated that a different state of affairs obtained than was the case with amount of supervision. As may be seen in Tables 35 and 35a a disproportional number of teachers who were appraised as above average-superior had experienced no difficulties in teaching, while a disproportional number of teachers appraised as average-below average had difficulties with discipline and classroom organization. The effects were stronger in pattern A systems than in pattern B systems.

TABLE 35. NUMBER OF TEACHERS IN PATTERN A SYSTEMS APPRAISED AS ABOVE AVERAGE-SUPERIOR OR AS AVERAGE-BELOW AVERAGE, BY TYPE OF TEACHING DIFFICULTY.

	Average-below average	Above average-superior
No difficulties	3	20
Misc. difficulties	18	11
Difficulties with discipline and organization	15	6

$X^2 = 18.089$, 2 df, $p < .005$

TABLE 35a. NUMBER OF TEACHERS IN PATTERN B SYSTEMS APPRAISED AS ABOVE AVERAGE-SUPERIOR OR AS AVERAGE-BELOW AVERAGE, BY TYPE OF TEACHING DIFFICULTY

	Average-below average	Above average-superior
No difficulties	1	22
Misc. difficulties	12	21
Difficulties with discipline or organization	7	5

X^2 = 12.612, 2 df, p .005

The data appearing in Tables 35 and 35a, especially those in Table 35, suggest that one component of supervisory appraisals of teacher success was the presence of difficulties in teaching, and especially difficulties with discipline and classroom organization. It is also true, as shown in the preceding section, that amount of supervision received was related to difficulties with discipline and organizational problems in pattern A systems. Amount of supervision and supervisory appraisals, however, were not related to each other. The apparent inconsistency in these relationships in type A systems is undoubtedly based in part on the fact that the amount of supervision given was obtained primarily from principals, while appraisals of success were obtained from supervisors. While each group appears to have based its actions or appraisals on the difficulties experienced by the teacher, the appraisals made by the supervisors encompass the success of the supervisory actions of the principal. Hence, in making appraisals, supervisors may legitimately take into account improvements made by teachers under the supervision of principals, and thus appraise the teacher as relatively successful in spite of the fact that the principal may have had to spend considerable supervisory time with the teacher and might not himself be disposed to consider the teacher especially successful for this reason.

Predictors of Appraisals of Teacher Success. The pre-test scores on the MTT and the TTR, and the TCS Factors were all potentially predictors of the supervisory appraisals made at either the end of the first year of experience (for the 1962 sample), or the end of the second year of experience (for the 1961 sample). To determine the extent to which these variables were predictors, teachers in pattern A systems were first separated from those in pattern B systems.

Within these groups, primary and intermediate teachers were separated for the analyses involving the MTT and TTR, but were pooled for analyses involving the TCS factors. The analyses were made via two group analysis of variance, with teachers appraised average-below average comprising one group ("low") and teachers appraised as above average-superior comprising the other ("high"). Raw Scores on the MTT, TTR and TCS Factors functioned successively as the continuous variables in the analyses.

Figures 2 and 3 show the profiles of mean scores, in Z score form*, for the high and low groups, for pattern A and pattern B systems respectively. In the first two rows of means in each profile, the results for intermediate and primary teachers are shown separately, in subsequent rows, intermediate and primary teachers are pooled.

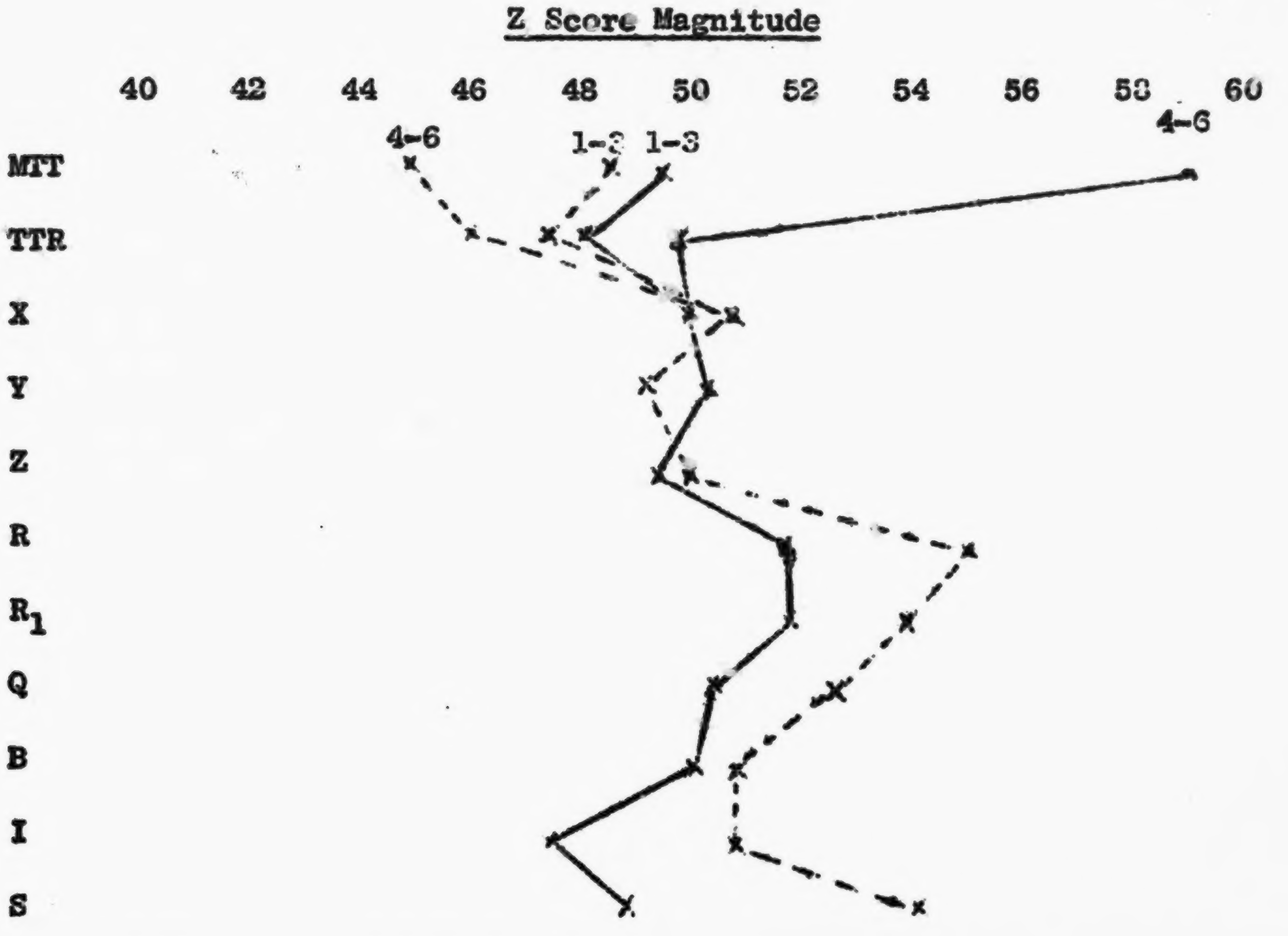

Figure 2. Differences in Mean Z Scores on MTT, TTR, and TCS Factors for "high" and "low" Groups, Pattern A Systems.

Key: ____________ high group, - - - - - - - low group

n = 38 high, 35 low

---

*Z scores were based on the distributions of each variable in the 1961 sample.

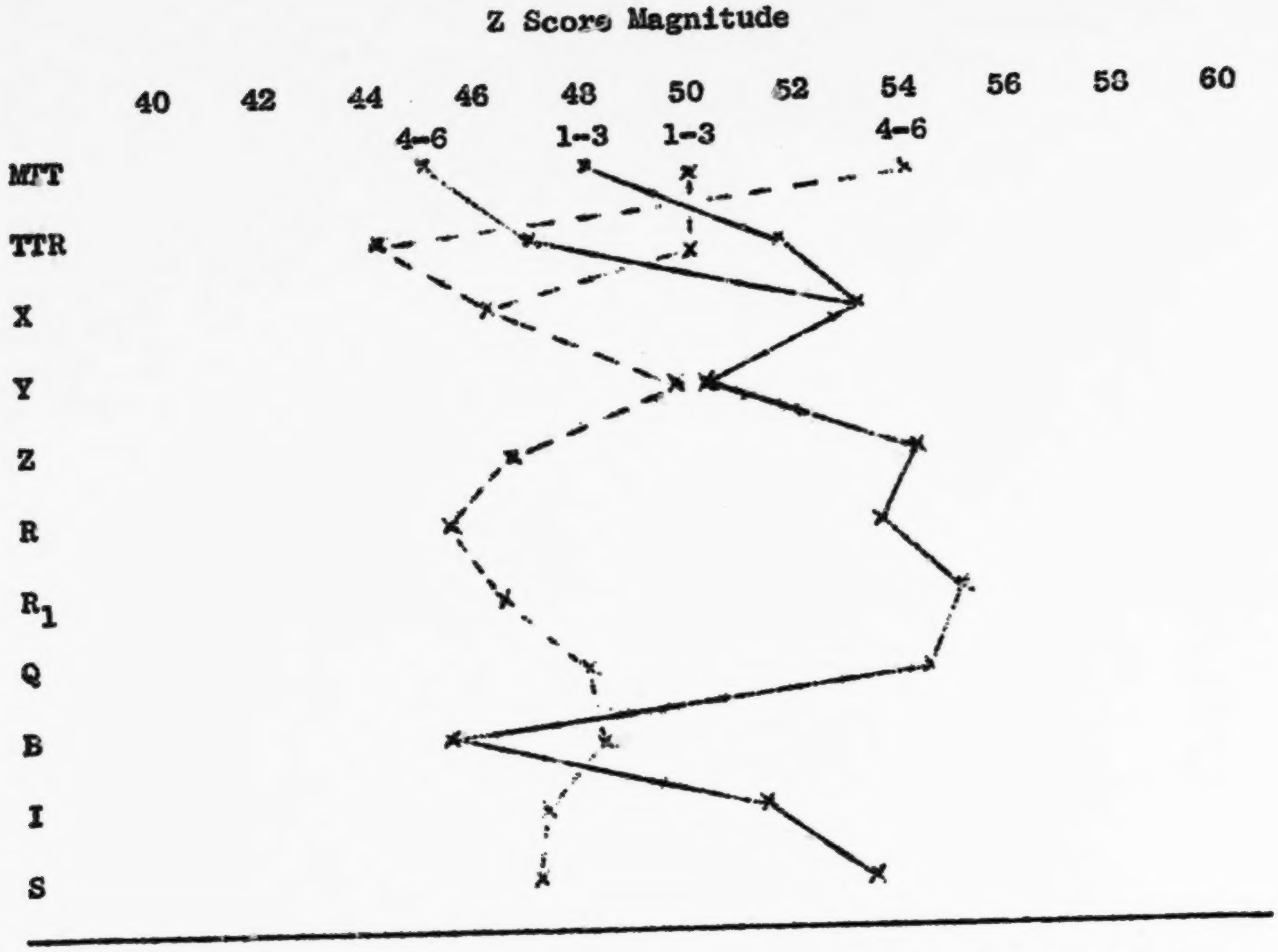

Figure 3. Differences in Mean Z Scores on MTT, TTR, and TCS Factors for "high" and "low" Groups, Pattern B systems.

Key: ____________ high group, - - - - - - low group

n = 45 high, 22 low.

Of the means arrayed in figure 2, only those for the high and low groups in grades 4-6 on the MTT differed significantly ($F = 18.55$, $p < .001$), with the high group greatly out performing the low group. Hence the MTT was a strong positive predictor of success among intermediate teachers in pattern A systems. In pattern B systems, as shown in Figure 3, however, the MTT is a negative predictor of success among intermediate teachers with the difference between high and low groups significant at $p < 05$ ($F = 5.53$) and the high group having the low mean. In addition teachers in the high and low groups in pattern B systems differed significantly on TCS Factors X (friendly, understanding) ($F = 9.13$, $p < 01$), Z (stimulating imaginative ($F = 5.74$, $p < 05$) R (favorable attitude toward pupils)

($F = 6.10$, $p < 05$), and R_1 (Favorable attitude toward democratic pupil practices) ($F = 6.03$, $p \leq 05$).

These findings distinctly suggest that quite different teacher characteristics were associated with success in different types of school systems. Pattern B systems appear to have emphasized personal-social variables in success, while pattern A system stressed task performance, especially in arithmetic. These particular differences in emphasis, together with the fact that pattern A and pattern B systems differ in the ratio of middle class to lower class schools, suggested the possibility that the differences in emphasis between the system types center in differences in the prevailing type of child with which the teachers on each system had to work.

The latter possibility was investigated, insofar as it could be investigated in the present study, by maintaining separation between pattern A and pattern B systems, dividing teachers within these system types according to whether they taught in working class or middle class schools, then observing differences between the average-below average group and the above average-superior group on the TCS factors. The results by system type may be observed in Figures 4 and 5.

As observation of the profiles in Figure 4 suggests, separation of teachers in working class and middle class schools in pattern A systems did not modify the relationships between the high and low groups on the TCS Factors. These factors remain non-predictive of success in this type of system irrespective of the type of school in which the teacher teaches. In pattern B systems, on the other hand, separation of teachers in middle class and working class schools resulted in the enhancement of differences between high and low groups in middle class schools, but a decrease in the differences between these groups in working class schools. Of the differences shown under "Middle Class Schools: in figure 5, those for Factor X differ at $p < .005$ ($F = 11.30$) for Factors Z at $p < 05$ ($F = 5.09$) for Factor R at $p < .025$ ($F = 6.56$) for Factor R_1 at about $p = 06$ ($F = 3.99$) for Factor I at $p < .025$ ($F = 5.63$) and for Factor S at $p < .025$ ($F = 6.08$). None of the differences between means shown under "Working Class Schools" in Figure 5, however, were significant, although the extremely small number of teachers with TCS scores teaching in working class schools in pattern B systems necessitates regarding this outcome as highly tentative.

An important factor underlying, but not directly observable in, the results shown in Figure 5 is that 80% of the teachers teaching in working class schools in pattern B systems were drawn from only two of the seven systems showing this pattern, viz., systems C and I. Each of these systems deviates from the most typical pattern for systems of this type in having a relatively high ratio of beginning

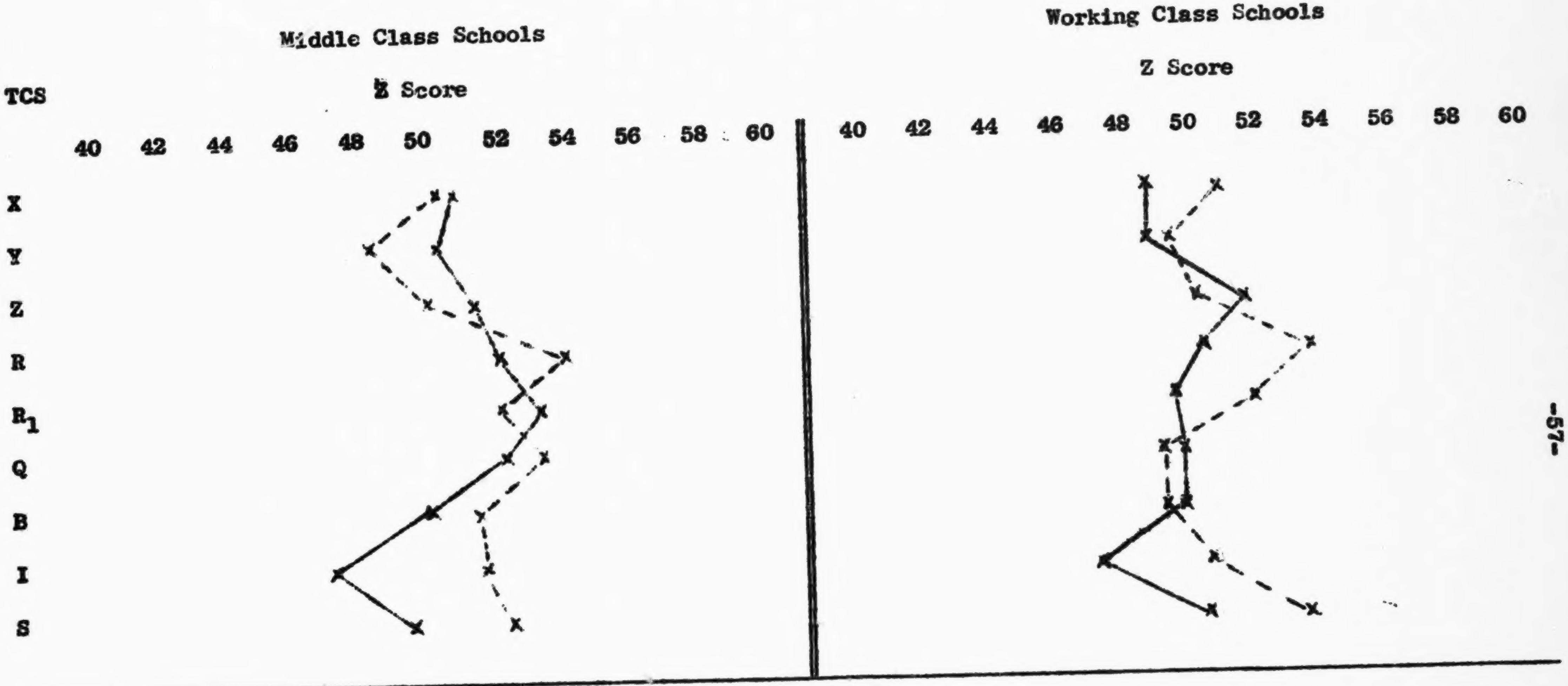

Figure 4. Differences in Mean Z Scores on TCS Factors for "high" and "low" Groups Teaching in Working Class or Middle Class Schools in Pattern A Systems.

Key: __________ high group, - - - - - - - low group

n for middle class = 19 high, 19 low, n for working class = 16 high, 19 low

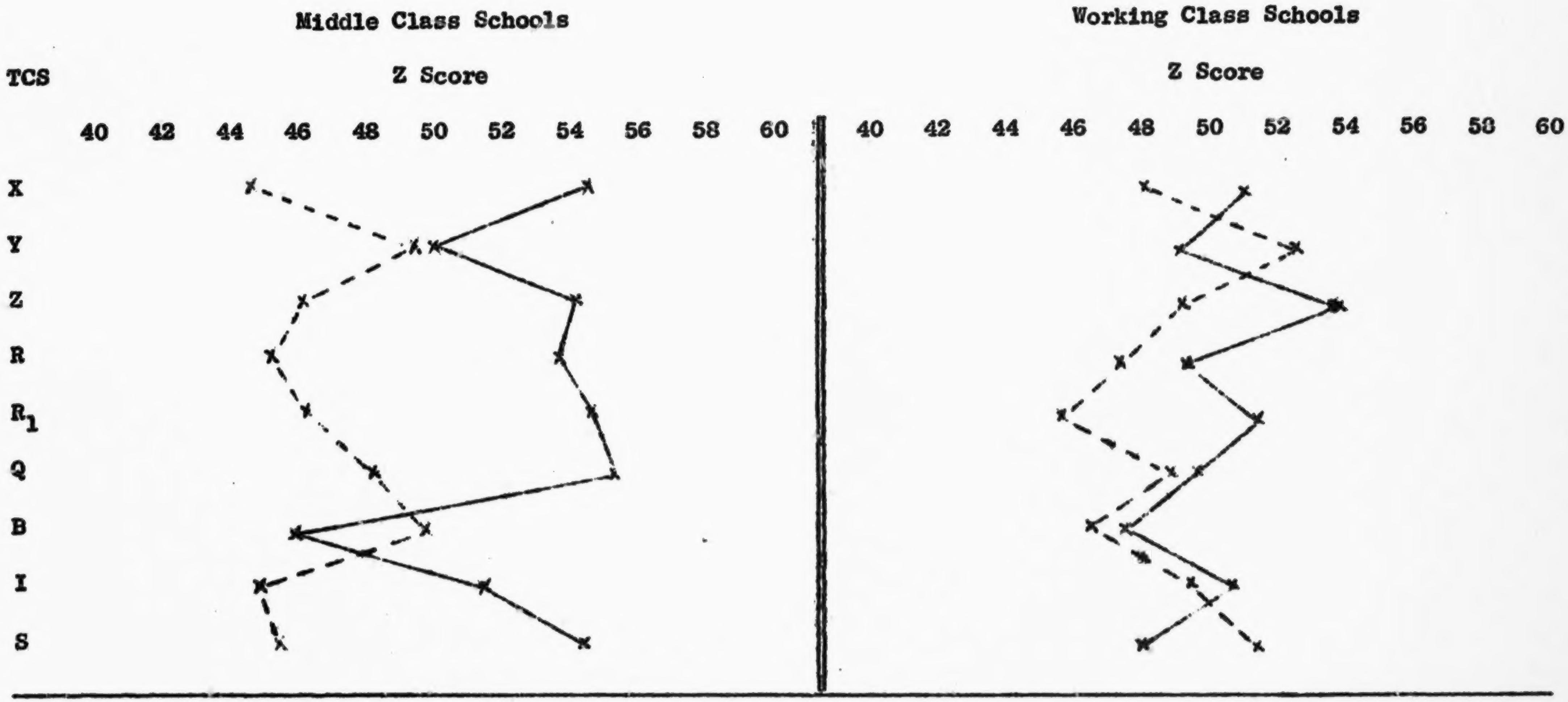

Figure 5. Differences in Mean Z Scores on TCS Factors for "high" and "low" Groups Teaching in Working Class and Middle Class Schools in Pattern B Systems.

Key: ___________ high group, - - - - - - - low group

n for middle class = 32 high, 16 low, n for working class = 13 high, 6 low.

teachers in working class schools.* This deviation, it might be noted, is in the direction of pattern A systems. In turn, the deviation of these systems appears to be accompanied by the loss of predictive power by the TCS scales. This is to say, that as the proportion of beginning teachers in working class schools in the school system increases, there appears to be a corresponding decrease in the predictive power of the TCS scales with respect to appraisals of teacher success.

Now the general preliminary interpretation to be made of the data shown in Figures 4 and 5 is not that the TCS scales are predictive of success for teachers in middle class schools but not predictive of success for teachers in working class schools; this interpretation is contradicted by the data in Figure 4. Rather, the preliminary interpretation is that the controlling factor in whether the TCS scales are predictive is the ratio of teachers in working class schools to teachers in middle class schools within a school system, and by implication, the ratio of working class to middle class children in that system. When the ratio of working to middle class schools stands at about 1:1 or less, the TCS scales appear to be predominantly irrelevant to teaching success as appraised by supervisors; but as the ratio increases toward 1:2 or more, the TCS scales appear to become increasingly effective predictors.

In making this preliminary interpretation it is especially important to note that both the type of school system and the type of school within the system are involved in the associations between the personal-social characteristics of the teacher and appraisals of teacher success. It is not just teaching in middle class schools which leads to the association, but rather, teaching in middle class schools in a system in which supervisory personnel are appraising success in an institutional context characterized by a prevalence of such schools. Correspondingly it is a change in this context towards proportionally fewer middle class schools which leads to a modification of the associations, not the absolute presence or absence of teachers teaching in middle class or working class schools.

Variables Associated With Teacher Mobility

Teacher mobility - non mobility was important to the study in three ways. First, the results previously reported were based on the inclusion of all teachers for whom data were available. Some of these teachers subsequently left their positions, however, and it is relevant to ascertain whether those who left differed from those who stayed with respect to the characteristics measured. Second, mobility has a certain limited utility as a criterion of teaching success. It is limited because teachers leave their initial positions for a wide variety of reasons, including marriage and pregnancy, which cannot be wholly construed as indicators of failure in teaching, nonetheless both marriage and pregnancy in certain cases may perhaps be interpreted to be instrumental in escaping from teaching. Third, the removal of teachers who were mobile after the first or second year of experience made it possible to determine the extent to which the residual (or "staying") groups differed from each other by grade level and system type, thus supplementing results in earlier sections.

---

*See Table 7 for the number of beginning teachers in working class and middle class schools in each system.

Variables Differentiating Mobile and Non-Mobile Teachers. To determine whether mobile as opposed to non-mobile teachers differed on either the MTT, the TTR, or the TCS Factors, grade level groups were separated, then mean differences on each variable between those who stayed and those who left by system type were observed via analysis of variance. In Figure 6 are shown the profiles of means for mobile and non-mobile teachers in grade 4-6, by system type.

Of the differences shown in Figure 6, only those for Factor Y (business like, responsible) were statistically significant ($F = 5.29$, $p < .025$), with those who left having the lower Y scores in both pattern A and pattern B systems. The MTT in pattern A systems, and Factor Q in both pattern A and pattern B systems were not quite significant at $p = .05$, although they appear to be different in the profiles shown. The fact that they appear to be different, but are not, is attributable to variations in the variances of the groups involved compared to the variance of the 1961 sample on which the Z scores norms are based. This factor also underlies the visually smaller but nonetheless statistically significant differences between means on Factor Y.

Differences between mobile and non-mobile primary teachers by system type are shown in Figure 7. Of the differences between groups shown in this figure, only that for Factor X in pattern A systems was statistically significant ($F = 4.51$, $p < .05$).

As Figures 6 and 7 suggest, there were relatively few differences between mobile and non-mobile elementary teachers. While Factor Y and to a less extent the MTT and Factor Q have some power to discriminate between the two groups among intermediate teachers, and Factor X some discriminatory power among primary teachers in pattern A systems, none of these variables shows sufficient strength to be of much practical use in predicting who will leave and who will stay in their initial teaching positions.

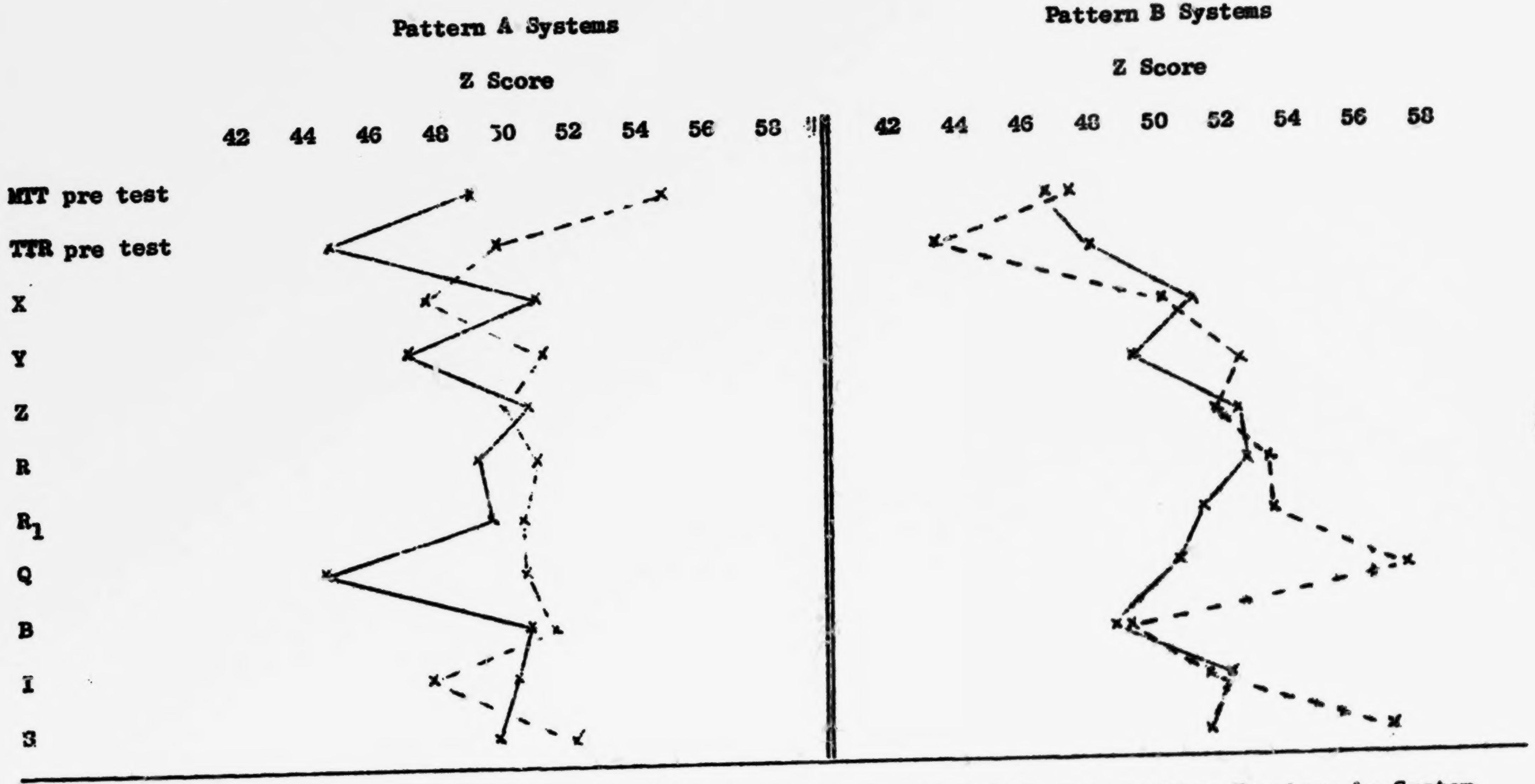

Figure 6. Mean Z Score on MTT, TTR and TCS Factors for Mobile and Non-Mobile Intermediate Teachers by System Type.

Key ____________ Mobile,- - - - - - - Non-Mobile

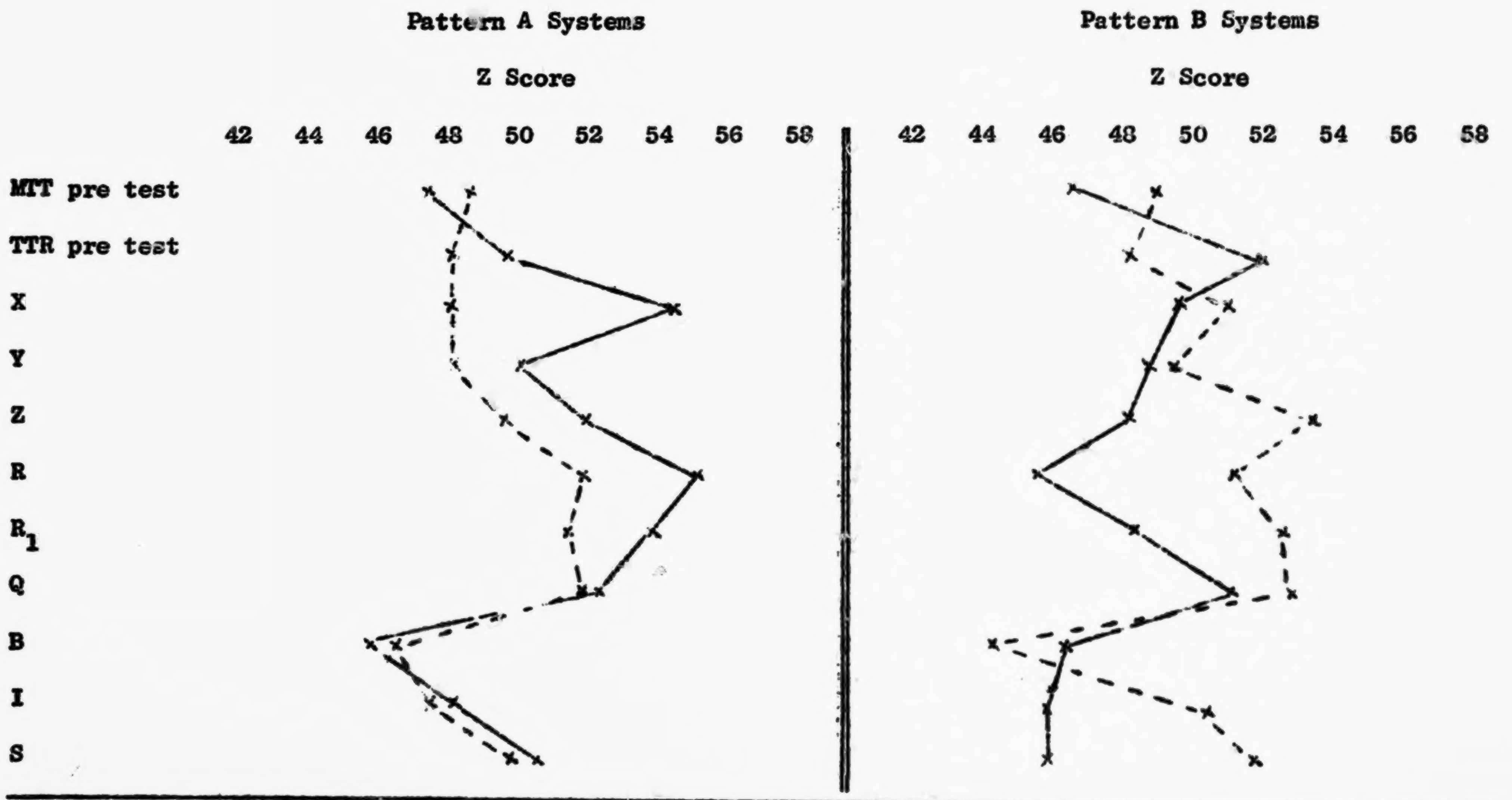

Figure 7. Mean Z Scores on MTT, TTR and TCS Factors for Mobile and Non-Mobile Primary Teachers by System Type.

Key ________ Mobile, - - - - - - - Non-Mobile.

Variables Differentiating Non-Mobile Teachers by Grade Level and System Type.

The examination of the MTT, TTR and TCS scores of non-mobile teachers by grade level and system type was conducted not only to supplement the findings reported in previous sections, but also to show how the MTT, TTR and TCS variables would have been related to grade level and system type had the study utilized cross sectional sampling in the fall of 1963, rather than longitudinal sampling, beginning in 1961, In a slightly different perspective, the examination of the terminal status of non-mobile teachers on the MTT, TTR and TCS in the general context of the longitudinal aspects of the study permits one to gain some understanding of how it was that the terminal status of the various grade level and system groups came about.

For the analyses, all teachers teaching as of the fall of 1963 in the system in which they were initially employed were used. This group was then divided by system type and redivided according to grade level, 1-3 and 4-6, and the analysis conducted via two way analysis of variance with the MTT, TTR post tests and the TCS functioning successively as the dependent variables. The means underlying the analyses may be observed in Figure 8.

Of the total array of variables in Figure 8, two were significant. First, on the MTT post test, grades 4-6 teachers performed significantly higher than grades 1-3 teachers ($F = 8.31$, $p < .01$) and teachers in Pattern A systems performed significantly higher than teachers in pattern B systems ($F = 4.92$, $p < .05$). This particular pattern of results was traceable to the high performance of the intermediate teachers in pattern A systems, which carried both main effects in the two way analysis. The high terminal performance of this group, as shown in earlier sections, is predominantly to be accounted for by the differential effects of the supervision received by the group compared to the remaining groups shown in Figure 8. Thus, differences between grade level groups and between system types in performance status appear to occur primarily through differential learning rather than through selection processes. It is of some interest to note in this respect, that while the pre-test performance of the total group of intermediate teachers who left pattern A systems was somewhat lower than those who stayed, as shown in Figure 7, the post-test performance of the eight teachers who left after taking the post-test was approximately five Z score units higher than those who stayed. Thus mobility served to decrease the differences in terminal status between groups rather than increase it, and may be viewed as working somewhat in opposition to the effects of supervision, at least in the present instance.

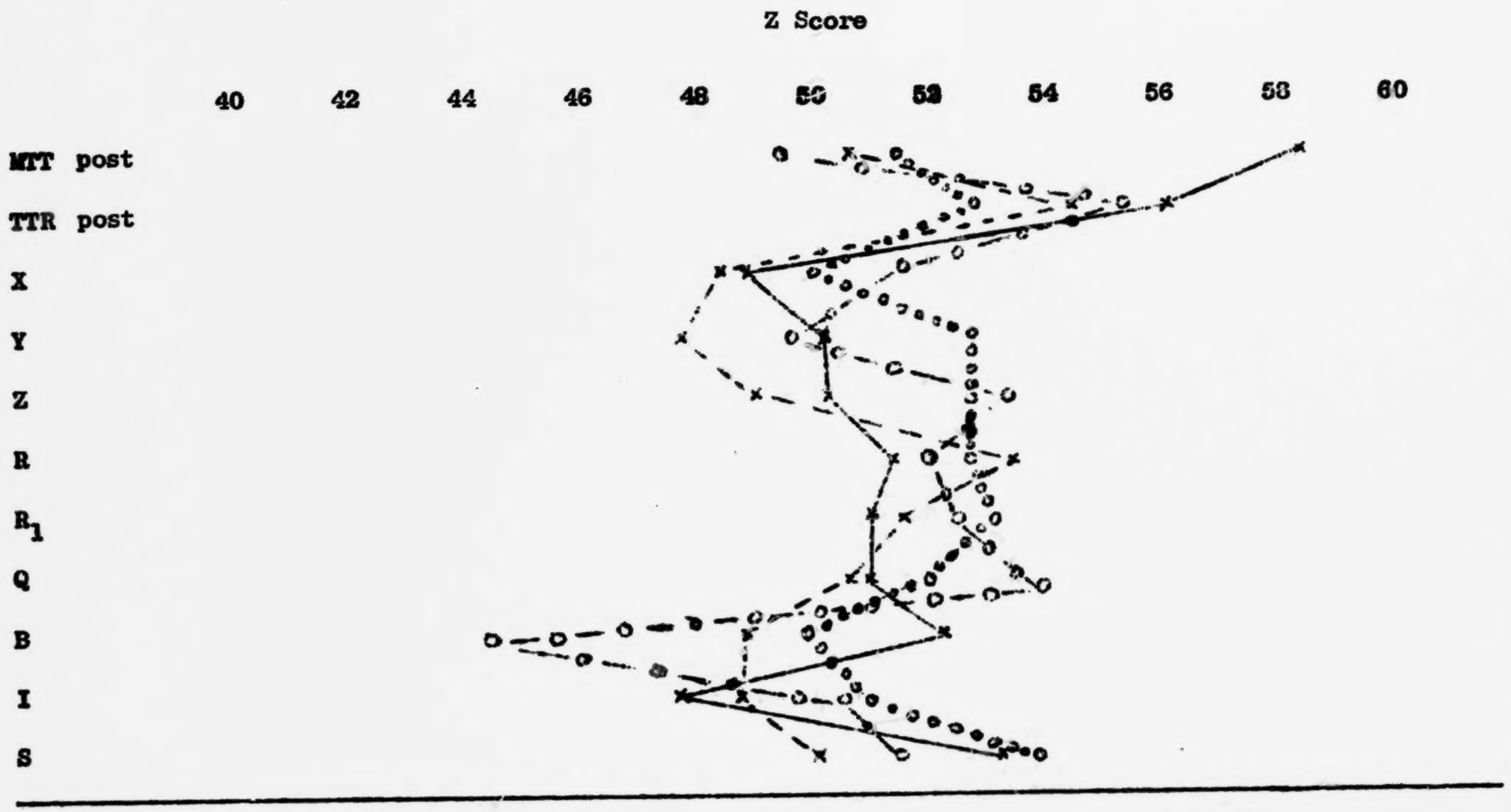

Figure 8. Mean Z Scores for MTT, TTR and TCS Factors for Non-Mobile Teachers by Grade Level and Type.

Key: Grades 4-6, Pattern A Systems ________
Grades 1-3, Pattern A Systems - - - - -
Grades 4-6, Pattern B Systems oooooooo
Grades 1-3, Pattern B Systems o-o-o-o-

The second variable yielding significant differences between groups as shown in Figure 8 was Factor B from the TCS. The effects of this factor were significant both by grade level ($F = 4.54$, $p < 05$) and by system type ($F = 3.93$, $p = 06$). These effects appear to have occurred through the slightly more traditional teachers leaving grades 1-3 in pattern B systems, while the slightly more permissive teachers left grades 4-6 in pattern A systems, thus bringing about sharper differences between grade level and system types on Factor B than was the case when the teachers were initially employed.

With the exception of the two variables noted above, teachers in the primary and intermediate grades in pattern A and pattern B systems are quite similar in the degree to which they possess the characteristics measured, at the end of the first or second year of experience. As previously shown, the differences in teacher characteristics between system types appear to lie primarily in the way supervisory personnel value these characteristics in appraising success, rather than in the degree to which teachers in the various systems possess these characteristics. The latter statement is based on results obtained using both mobile and non-mobile teachers, but it remains true if only non-mobile teachers are utilized in the analyses, as may be observed in Figure 9.

Two features of the data arrayed in Figure 9 are of particular interest. First, Factors X ($p < .01$), Z ($p < .025$) R ($p < .05$) and I ($p < .05$) significantly distinguish between non-mobile teachers appraised as average-below average and above average-superior in pattern B systems, but not in pattern A systems. Thus in spite of the fact that non-mobile teachers in these two types of systems do not differ in the degree to which they possess these characteristics, as shown in Figure 8, the value placed on the characteristics in the supervisory appraisals in the two types of systems is quite different. This result clearly demonstrates that the determining factor in whether a beginning teacher is appraised as relatively successful or relatively less successful does not in fact depend on the degree to which he or she possess a given characteristic or set of characteristics, but rather on the way in which the level of each characteristic possessed is valued by the school system in which he or she is employed. It follows that statements about teaching success, on the criterion of supervisory appraisal, are not statements about a property of the teacher issuing from his various characteristics, but are rather statements about the relationship between the characteristics of a teacher and characteristics of situation in which he teaches.

A second feature of interest in Figure 9 is the reduction in the degree to which the MTT post-test distinguishes between average-below average as opposed to above average-superior intermediate teachers in pattern A systems, compared to the degree to which the

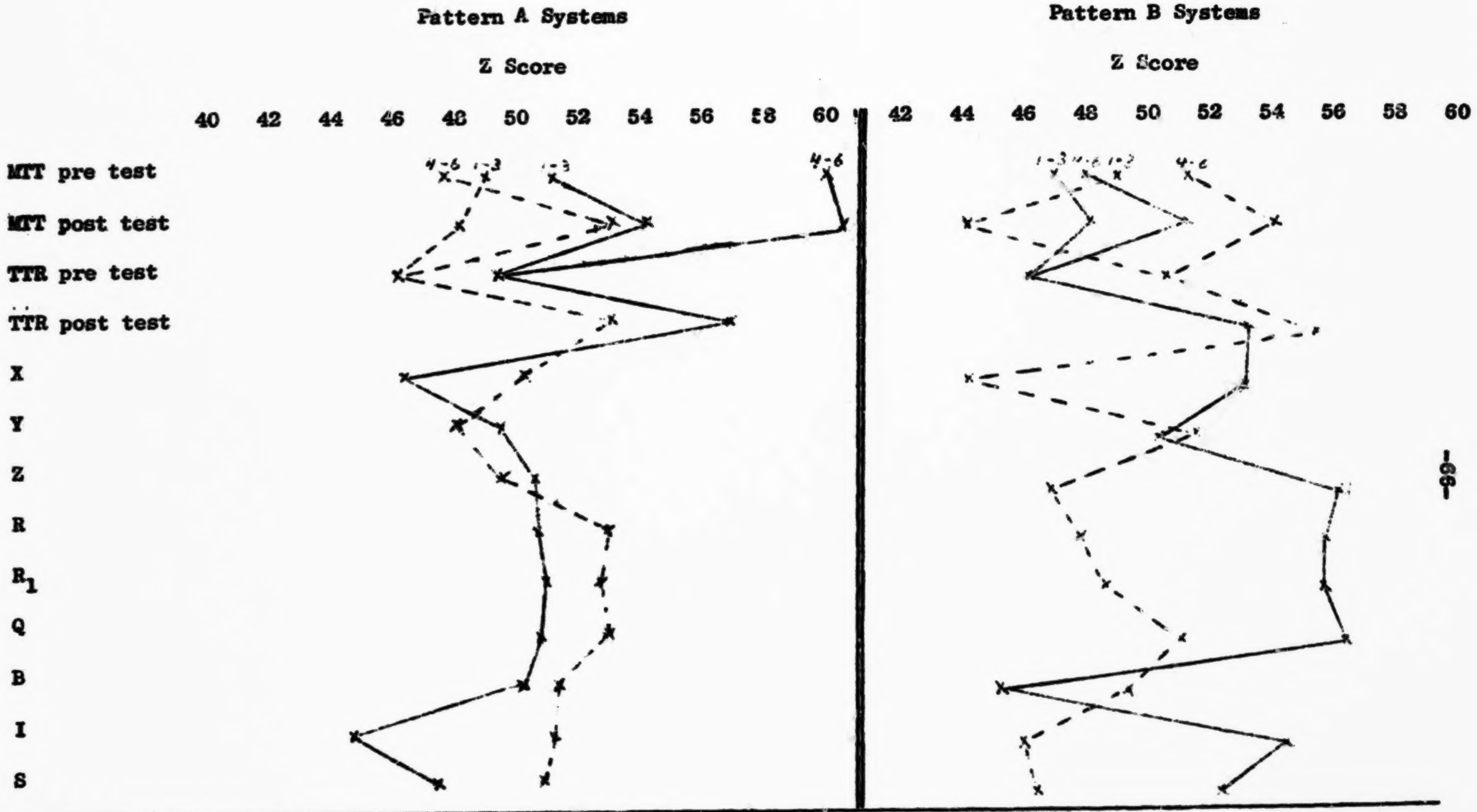

Figure 9.
Mean Z Scores for Non-Mobile Teachers Appraised as Average-Below Average or Above Average-Superior by System Type.

Key: ____________ above average-superior - - - - - - average-below average

MTT pre-test distinguishes between these two groups. The F value for the differences between groups on the MTT pre-test was 9.26, $p < .01$, while on the post-test the F value was 4.56, $p < .05$. Now the significance of this reduction has two aspects. First, it demonstrates the difference between the predictive validity and the concurrent validity of a test. The MTT clearly is stronger as a predictor of supervisory appraisals than it is as a concurrent variable with respect to these appraisals. This state of affairs is distinctly disturbing from the viewpoint of research strategy, since it suggests that cross-sectional sampling methods and concommitant examination of the concurrent relationships between independent and dependent variables could lead to the dismissal of variables which show poor concurrent relationships, but which might have shown excellent predictive relationships, had longitudinal methods been used. Second, the reduction in the power the MTT from pre to post test in distinguishing between groups appraised as having the greater or the less success appears to be directly traceable to the supervisory activity of the school systems in which the teachers are employed. Supervision in these systems has the effect of making beginning intermediate teachers increasingly homogeneous in MTT performance, that is, the teachers who were initially low in performance increase their performance under supervision at a rate faster than those initially high in performance. Under these conditions, if supervisory appraisals in fact reflected only the current classroom performance of the teacher, a point should eventually be reached at which it would be virtually impossible to find a variable having concurrent validity with respect to supervisory appraisals, since both teacher performance and the supervisory perception of teacher success would be extremely homogeneous.

Interrelationships Among Task Performance and Personal-Social Variables

Three principal questions were involved in examing the interrelationships among task performance and personal-social variables. First, does a direct relationship exist between changes in performance on the MTT and TTR and the personal social variables measured by the TCS? Second, what is the factorial composition of the MTT, TTR and TCS, and third, does this composition vary by grade level?

Relationships Between Changes in Task Performance and Variables from the TCS. To ascertain these relationships, teachers of grades 1-3 were separated from those in grades 4-6. The change scores for each task for each group were then correlated with the TCS scales via product moment correlation. The resulting matrices may be observed in Tables 36 and 37, for teachers of grades 4-6 and grades 1-3 respectively.

TABLE 36. CORRELATIONS AMONG CHANGE SCORES ON MTT AND TTR TASKS, AND TCS SCALE SCORES FOR INTERMEDIATE TEACHERS.

	M T T					T T R				
		Tasks					Tasks			
T C S Scale	Sum of Tasks	1	2	3	4	Sum of Tasks	1	2	3	4
X	07	04	13	08	14	-14	-04	-30*	07	02
Y	-12	03	06	16	-21	-10	-07	13	-02	-15
Z	-21	-18	29	12	-31*	-25	05	-47*	06	09
R	-16	-10	02	15	-20	-27	01	-17	-16	20
R_1	-09	01	10	09	09	-18	-06	-06	-14	05
Q	12	-03	-13	-02	-07	-16	-12	-03	-07	-04
B	17	21	09	03	12	06	07	20	07	02
I	-01	14	12	01	-01	03	09	10	04	05
S	13	11	08	25	21	-23	04	-20	-2	30*

*p = 05 or less

In both tables 36 and 37, the number of correlations significant at p = 05 is within what would be expected by chance. Moreover, there is very little tendency for associations statistically significant among intermediate teachers to be also significant among primary teachers. It therefore seems that the best interpretation to be given to the data appearing in these tables is that they do not support the notion that there are significant direct relationships between any TCS Factor, and changes in task performance on either the MTT or TTR. Rather, as suggested earlier on the basis of a somewhat different type of analysis, the relationships between changes in task performance and TCS Scales are indirect or moderated relationships.

TABLE 37. CORRELATIONS AMONG CHANGE SCORES ON MTT AND TTR TASKS, AND TCS TASKS, AND TCS SCALE SCORES FOR PRIMARY TEACHERS.

	M T T					T T R				
T C S Scale	Sum of Tasks	Tasks 1	2	3	4	Sum of Tasks	Tasks 1	2	3	4
X	-01	-15	29*	13	16	-02	04	-10	-13	04
Y	19	05	35*	21	-03	-05	-06	02	-03	-03
Z	-12	-13	02	19	-20	-09	-04	-21	-17	00
R	09	-10	19	21	-07	-04	09	-27*	-16	05
R_1	29*	11	32*	10	13	-02	-13	-19	-15	08
Q	05	03	18	17	-16	-16	-08	-26*	-14	-00
B	00	03	-04	-03	06	-01	-07	07	10	-05
I	-13	00	07	03	-20	-06	-00	-18	-06	02
S	03	-01	19	18	-15	-07	05	-25*	-11	-01

*p = 05 or less

Factor Composition of the MTT, TTR and TCS. The factor composition of the MTT, TTR and TCS was examined via BMED Program 17, utilizing the principal components method with varimax rotation to orthogonol factors. Four analyses were made, two for intermediate teachers, and two for primary teachers. The two analyses for each grade level group differ from each other only in that the first analysis in each case included the total scores for the MTT and TTR pretests, while in the second analyses, these scores were omitted. The first analysis for each group was of interest primarily in determining relationship among the TCS Scales and MTT and TTR Total Scores when the contribution of each task to the MTT and TTR totals was left in the matrix. In general these two analyses show how each task fed into the total score of its respective test, and how test totals relate to each other, however, the factoral composition of the tasks, independent of total score, is distorted in these analyses, and would lead to spurious interpretations of factorial composition within the MTT and TTR, since the correlation of each task with its own contribution to the total score for its test becomes a determinant of the factors extracted.

The distortion of factors within the MTT and TTR is corrected in the second set of analyses, in which the total scores for MTT and TTR were omitted. These analyses consequently permit a more accurate interpretation of the factorial composition of the MTT and TTR, but do not show how total scores between tests are related. The latter must be obtained from the first set of analyses.

Tables 38 and 39 give the intercorrelations among the variables analyzed among intermediate and primary teachers respectively. For the first set of analyses, all the correlations shown for each group entered the analysis for that group. For the second analysis, however, the correlations appearing in row 1 and row 5 of each matrix did not enter the analysis for the group.

Observation of the correlations between the TCS Scales and the task and total scores on the MTT and TTR, as shown in Tables 38 and 39, suggest what was subsequently revealed in the first set of factor analyses, as shown in Table 40, viz., that the TCS Scales, the TTR totals and the MTT totals load on orthogonal factors (factors 1, 2 and 3 respectively for both intermediate and primary teachers, as shown in Table 40). Observation of those aspects of Tables 38 and 39 which show the correlations between tasks and between tasks and total scores for the MTT and TTR respectively also indicate what subsequently appeared in the factor analyses. Among intermediate teachers (table 38) the correlations among MTT tasks 2, 3 and 4 and MTT total are positive and moderately high, which is primarily an indication of the contributions these tasks made to the total score, while the intercorrelations among these tasks is positive but small. These tasks correspondingly show relatively high loadings on factor 3 among intermediate teachers in Table 40, the factor identified with the MTT total. Among primary teachers, a similar state of affairs holds with respect to factor 3, although the low intercorrelations among tasks 2, 3 and 4 on the MTT, and the moderate to high correlations between MTT task scores and total scores indicates with greater clarity than among intermediate teachers the immense influence the correlations between tasks and total score had in determining factor 3.

Factor 2 in Table 40 (for both primary and intermediate teachers) like factor 3, had its determinants in the task-total score correlations on the TTR. In this instance the high correlations between TTR total and TTR tasks 1 and 3, and the moderate correlations between task 1 and 3 largely determined the factor, especially among intermediate teachers. Among primary teachers the moderate positive correlations between task 2 and tasks 1 and 3 also contributed to the factor.

The principal significance of the first two factor analyses, as briefly described above, is, first, that TCS, MTT and TTR scores

TABLE 38. INTERCORRELATIONS AMONG MTT PRE-TEST, TTR PRE-TEST AND TCS VARIABLES FOR INTERMEDIATE TEACHERS.

	MTT Tasks				TTR Tasks					TCS								
	1	2	3	4	Total	1	2	3	4	X	Y	Z	R	R_1	Q	B	I	S
MTT Total Score	43	49	49	86	24	24	14	05	11	-11	04	-06	-10	-09	01	-05	-16	-06
Task 1		08	-09	21	30	20	16	20	19	-07	-04	-05	-13	-26	-13	-09	-27	-13
Task 2			43	14	32	32	14	20	02	-13	03	-17	-14	00	08	08	-01	-16
Task 3				16	28	32	19	11	-03	-08	-19	-05	-05	00	01	-11	02	-09
Task 4					-01	01	00	-12	09	-05	15	01	-02	-01	03	-02	-12	06
TTR Total Score						79	34	62	47	04	-02	-01	05	06	06	06	07	13
Task 1							03	36	14	06	09	-02	02	-04	12	-04	10	-06
Task 2								12	03	19	-23	24	03	-03	-04	-03	16	01
Task 3									-04	-02	03	12	10	05	10	-11	16	15
Task 4										-10	-09	00	-04	-17	-13	02	-23	-12
TCS Scales																		
X											33	73	71	55	59	-47	34	62
Y												35	41	38	53	-22	04	44
Z													67	41	52	-53	44	67
R														79	76	-45	44	83
R_1															76	-36	33	75
Q																-27	43	66
B																	-21	-35
I																		43

p at .05 = .25

n = 57

TABLE 39. INTERCORRELATIONS AMONG MTT PRE-TEST, TTR PRE-TEST AND TCS VARIABLES FOR PRIMARY TEACHERS.

	MTT Tasks				TTR Tasks					TCS								
	1	2	3	4	Total	1	2	3	4	X	Y	Z	R	R_1	Q	B	I	S
MTT Total Score	18	25	42	85	12	16	-06	15	-02	07	08	11	-05	-10	-04	09	18	04
Task 1		03	-01	-11	16	26	-06	06	01	09	16	08	08	02	13	09	08	13
Task 2			01	04	06	11	-08	05	-01	-12	-12	03	-00	-11	-16	01	06	-03
Task 3				08	03	03	09	07	07	-16	-02	-18	-19	-09	-05	11	07	-14
Task 4					06	06	07	11	00	15	09	16	-01	-05	08	03	13	07
TTR Total Score						80	43	70	51	02	02	00	-10	02	07	13	18	00
Task 1							16	38	16	03	00	-03	-13	03	00	13	12	-09
Task 2								42	-07	10	-07	12	05	09	11	07	17	09
Task 3									12	07	03	08	-05	03	00	05	12	03
Task 4										-10	13	-10	-06	03	11	05	07	06
TCS Scales											42	75	72	64	58	-42	44	65
X												31	36	45	52	-22	03	42
Y													68	53	43	-57	43	60
Z														77	68	-54	35	80
R															75	-56	30	70
R_1																-33	26	66
Q																	-14	39
B																		44
I																		

p at 05 = .21

n =

TABLE 40. FACTORS AND FACTOR LOADINGS ASSOCIATED WITH THE MTT, TTR AND TCS FOR PRIMARY AND INTERMEDIATE TEACHERS.

Variable	Intermediate Teachers--Factors						Primary Teachers--Factors					
	1	2	3	4	5	6	1	2	3	4	5	6
MTT Total Score	05	-14	96	17	06	04	02	07	-96	03	17	15
Task 1	12	-23	25	-79	09	09	07	-06	02	05	02	-91
Task 2	12	-46	53	29	03	11	03	-02	-09	-01	79	06
Task 3	05	-30	54	45	-35	07	20	-12	-46	34	-26	-11
Task 4	-04	17	84	-21	10	09	-07	-01	-90	-11	02	18
TTR Total Score	-03	-89	07	-09	-20	-36	03	-90	05	-39	04	16
Task 1	-04	-80	14	05	03	-18	08	-66	07	-24	19	-40
Task 2	-13	-13	08	-11	-80	-01	-10	-71	11	37	-24	-18
Task 3	-00	-78	-12	-23	02	32	-02	-79	-10	-03	02	05
Task 4	08	-17	02	-08	03	-92	04	-23	01	-83	-06	03
TCS Scales												
X	-82	01	-09	-08	-14	-02	-85	-07	-08	07	-05	-05
Y	-50	07	11	14	61	03	-47	14	-14	-32	-40	-29
Z	-83	13	-04	-10	-25	-12	-82	-06	-08	11	-20	04
R	-92	-05	-06	04	08	-01	-91	-08	09	03	01	-02
R_1	-79	-02	01	23	22	14	-34	-02	11	-09	-21	-00
Q	-80	-17	06	16	28	11	-74	-02	-05	-17	-37	-16
B	57	04	-05	29	21	08	64	-14	-11	04	-14	-21
I	-51	-15	-18	29	-27	28	-49	-29	-18	14	22	-02
S	-87	-16	02	08	11	01	-85	00	-02	-05	-07	-04

p at 05 = .21

n = 84

represent orthogonal factors, a fact which made possible the findings reported in earlier sections, in which these three main variables held quite different relationships to various criteria and between various groups. Second, with respect to the MTT and TTR total scores, it seems relatively clear that task 4 on the MTT and to a lesser extent, tasks 2 and 3 on this instruments were the controlling factors in the results obtained. On the TTR on the other hand tasks 1 and 3 were major contributors to the results obtained, with task 2 also contributing among primary teachers, but not among intermediate teachers

Other than observations of the orthogonality of the first three factors representing the TCS, the TTR and the MTT respectively and the way in which MTT and TTR tasks related to factors 2 and 3, no interpretation on the first set of factor analyses were made. The general interpretation of factors was reserved for the second set of analyses, for which the factors with eigen values larger than 1 and factor loadings are shown in Table 41.

As may be observed in Table 41, except for the first factor, the factors extracted from the primary and intermediate groups are somewhat different. Hence, the interpretation of the first factor is common to both groups, but the remaining factors must be interpreted separately.

Factor 1 is clearly identified with the TCS scales, especially scales X (friendly, understanding) Z (stimulating, imaginative), R (favorable attitude toward pupils), R_1 (favorable attitude toward democratic pupil practices), Q (favorable attitude toward school staff personnel) and S (emotional adjustment), and to a lesser extent with factors Y (responsible, businesslike), B (traditional vs. permissive) and I (verbal understanding). This factor thus appears to represent a favorable personal-social orientation, which includes not only relating well to others, but also elements of favorable emotional adjustment and ability to stimulate. While the relatively modest loadings of scales Y, B and I on this factor suggest that under oblique rather than orthogonal rotation, one or more additional factors might be extracted from the TCS, it seem nonetheless clear that most of the scales in this instrument measure a single factor, at least among beginning teachers. It might be noted in this respect that the differences between groups appraised as average-below average and above average-superior in pattern B systems tend to follow the factor loadings of the various scales on factor 1, while differences between groups having and not having discipline and organization problems in pattern A systems are associated with factors Y and B, which are not closely associated with factor 1. Similarly, factor Y discriminated between mobile and non-mobile intermediate teachers in pattern A and pattern B systems. In neither case, however, were those scales loading highly on factor 1 also discriminating. Thus one may reasonably infer that while scales Y and B share some commonality with the scales loading on factor 1, they also are sufficiently

TABLE 41. FACTORS AND FACTOR LOADINGS ASSOCIATED WITH MTT TASKS, TTR TASKS AND TCS SCALES FOR PRIMARY AND INTERMEDIATE TEACHERS.

Variable	Intermediate Teachers--Factors					Primary Teachers--Factors					
	1	2	3	4	5	1	2	3	4	5	6
MTT Task 1	05	03	77	-09	-27	08	05	84	-01	-12	07
MTT Task 2	11	-75	-00	03	-26	-03	04	07	-77	01	02
MTT Task 3	04	-79	-13	-28	-05	-21	-14	15	-26	44	-36
MTT Task 4	-08	-51	44	.18	39	07	03	-13	03	87	10
TTR Task 1	-05	-39	18	04	-64	-04	-47	53	22	12	28
TTR Task 2	-17	-16	12	-77	-05	09	-82	11	-18	-12	14
TTR Task 3	-01	-04	01	-03	-85	00	-77	08	09	12	18
TTR Task 4	06	01	58	-04	01	04	-08	02	-05	03	33
T C S											
X	-83	10	-01	-12	-02	84	-08	07	-03	14	07
Y	-50	-02	11	63	-05	48	20	-27	42	14	27
Z	-84	10	07	-23	12	82	-07	-03	21	11	-11
R	-90	05	-12	11	-06	91	05	01	-00	-10	-04
R_1	-76	-09	-30	25	01	85	-05	-02	-21	-12	09
Q	-78	-14	-19	31	-14	74	-02	15	-38	03	17
B	60	-02	-18	19	09	-64	-12	25	-12	14	01
I	-48	00	-49	-26	-21	47	-28	11	27	32	07
S	-86	04	-14	14	17	85	-.02	08	-06	02	06

independent of this factor to yield interesting discriminations not attributable to it.

Factor 2 among intermediate teachers is primarily identified with tasks 2 and 3 on the MTT, and to a lesser extent with task 4, MTT and task 1, TTR. The common factor underlying these tasks is interpreted to be closely aligned with ability to form concepts or categories. Tasks 2 and 3 MTT and task 1 TTR all require the teacher to form and interpret categories. Task 4, MTT, does not explicitly require category formation, but it does require very careful differentiation of categories according to the criterial attributes of the arithmetic examples involved in the task. This latter skill holds at least a logical relationship to category formation skills, thus the moderate loading of task 4 MTT on factor 2 seems quite reasonable.

Factor 3 among intermediate teachers is positively and closely associated with task 1, MTT and to a lesser extent with task 4 MTT and task 4 TTR. Task 1 MTT asks the teacher to relate arithmetic exercises to arithmetic goals, while task 4, TTR asks the teacher to relate or match reading exercises which could be used to accomplish the same objective in teaching reading. Task 4 MTT, on the other hand, requires the teacher to order problems into a teachable sequence according to difficulty level, which requires not only that he differentiate the categories to which the various problems belong, as mentioned above, but that he order the categories. While the skills involved in this factor do not seem to fit neatly into a known category of cognitive skills, they seem closely associated with the type of skills utilized in programming material for instruction, skills associated with relating and ordering or organizing materials. In addition to the tasks loading positively on this factor, Scale I on the TCS shows a moderate negative loading on it. This negative loading is associated with the negative correlations between task 1 MTT and Task 4 TTR and Scale I as shown in Table 38. The fact however, that this loading is negative is disturbing, because both the MTT and the reading tasks on which the TTR was based have been previously shown (11, 12, 14) to hold significant positive correlations with the A C E Psychological Examination, a test with distinct verbal components. Yet Scale I holds primarily negative associations with the variables comprising the MTT and TTR. Thus the interpretation given to the negative loadings of Scale I on factor 3 is not that this factor is negatively related to verbal intelligence, but that Scale I is not in fact a measure of intelligence, at least in the usual menaing of that word. In view of its relatively strong positive associations with the TCS variables loading on factor 1, it might better be construed as a weak measure of the teachers personal-social orientation.

Factor 4 among intermediate teachers is a bi-polar factor, defined on the one hand by task 2 TTR (negative) and Scale Y (positive) on the other. This factor is a minor one (eigen value 1.444) and its bi-polarity is based primarily on a small negative correlation (-23). Logically, task 2 TTR should load on factor 3, since it requires the teacher to order reading passages according to difficulty level and therefore contains a programming or ordering component. However, the independence of this task to factor 3 in addition to its polarity to Scale Y, makes factor 4 a virtually uninterpretable factor. Certainly there is no apparent reason why ability to discriminate and order the difficulty level of reading passages should be negatively related to responsible, businesslike teacher behavior.

Factor 5 among intermediate teachers is also a minor factor and is identified with task 1 and task 3, TTR. Each of these tasks is closely aligned with skill in recognizing and categorizing errors involving word attack skills among pupils. This factor thus involves both knowledge of the principal methods of word attack, and ability to categorize errors. The relation of this factor to factor 2 is of some interest because they to some degree clarify each other. Factor 2 was identified as a category or concept formation factor, and part of the variance of task 1 TTR is independently associated with this factor, which is clearly not specific to reading but involves a substantial arithmetic component. On the other hand, a moderate part of the variance of task 1, TTR is associated with factor 5, while a substantial part of the variance of task 3, TTR is associated with factor 5, but virtually none of its variance is associated with factor 2, in spite of the fact that category formation skills are involved in task 3. This state of affairs suggests that there is a tendency for category formation skills to be tied to content areas, with only a part of the skills identified with concept formation independent of a specific content area. Given the present tasks, a factor specifically identified with concept formation completely independent of content areas could not be extracted, however, it should be possible to construct tasks in the future from which the variance attributable to skill in concept formation can be sorted out from the variance involving knowledge of specific content areas.

The factor composition of the MTT and TTR among primary teachers was somewhat different than among intermediate teachers. The correlational basis of these differences may be observed by comparing the intertask correlations from the MTT and for the TTR for each group in Tables 38 and 39.

Factor 2 among primary teachers is most closely associated with tasks 2 and 3 TTR, and moderately associated with task 1 TTR. This factor is a relatively large one, and clearly resembles a combination of factors 4 and 5 among intermediate teachers. The factor does not appear to represent a particular set of cognitive skills, but seems more nearly to reflect knowledge of and ability to handle various aspects of teaching reading.

Factor 3 among primary teachers is closely associated with task 1, MTT and somewhat less closely associated with task 1 TTR. The loadings of these two tasks, each from different content areas, and apparently requiring somewhat different cognitive operations, on the same factor at first seems curious. Task 1, MTT, however, is the only task in the MTT which is associated at all with the primary grades, while task 1, TTR is also closely associated with these grades. Thus the commonality between these tasks, which from a correlational viewpoint is not large, but nonetheless present, appears to stem from the familiarity primary teachers would have with such tasks, rather than from a common cognitive operation or a common content area. This factor thus introduces an essentially new dimension into the factor analyses conducted, viz., commonalities attributable to grade level at which the tasks occur.

Factor 4 among primary teachers, like factor 4 among intermediate teachers, is a minor bi-polar factor that is essentially uninterpretable. The negative aspect of the factor is clearly associated with Task 2, MTT, while the positive aspect rests primarily on a non-significant negative association (-.12) between task 1 and Scale Y. How one might interpret these associations other than as spurious[*], is not apparent from examination of the components involved.

Factor 5 among primary teachers is centrally associated with task 4 MTT, and to a lesser degree with task 3, MTT. The commonality of these two tasks for primary teachers probably lies in their content, rather than in common cognitive operations, since both task 2 MTT and task 1, TTR would be expected to load on the factor if concept formation as a cognitive operation were focally involved.

Factor 6 among primary teachers is closely associated with task 4, TTR, and is thus identified with ability to match reading exercises which lead to the same reading objective from the viewpoint of cognitive operations, task 4 TTR should load on the same factor as task 1, MTT. Its failure to do so among primary teachers is undoubtedly attributable to the differences in content between these two tasks which appears to obliterate the commonalities between the tasks in this group.

Differences in Factors Among Primary and Intermediate Teachers. The apparent differences in factors other than the first factor extracted for primary and intermediate groups may simply reflect unreliabilities in tasks in the MTT and TTR plus the fact that the two factor analyses are based on only 34 and 57 subjects respectively. There are, on the other hand, logical grounds for expecting differences in the factors extracted for these groups. First, the MTT tasks have never been shown to be valid for primary teachers. Correspondingly they represent an essentially irrelevant performance domain for this group. Second, and to some degree as a reflection of the first point, the total performance of intermediate teachers on the MTT and TTR was considerably more homogeneous than for primary teachers, not only from the viewpoint of the intercorrelations among the tasks, but also from the viewpoint of the mean performance on them. To put the matter another way, intermediate teachers show a generally better integration of task performance skills than do primary teachers, which in view of the fact that the tasks were constructed primarily for intermediate teachers, is not suprising.

---

[*] The parallel form reliability of task 2 MTT is very low ($r = .05$) which makes interpreting a factor based primarily on this task extremely hazardous.

The better integration of performance skills among intermediate teachers appears in turn not only to be associated with the emergence of fewer factors but also with a more coherent pattern of factor loadings on each factor. The heterogeneity of performance among primary teachers, especially on the MTT, on the other hand, is associated with a dispersion of factors and rather odd patterns of factor loadings, approximately what one would expect from a group performing work tasks irrelevant to their work. It may be noted in this respect that the factor loadings on factor 1, representing the TCS, which is relevant to both primary and intermediate teachers, are almost identical for each of these groups.

CHAPTER IV

HYPOTHESIS GENERATION AND SYNTHESIS OF RESULTS

In the preceding chapter preliminary interpretations of the various results of the study were made. The purpose of this chapter is to generate a network of hypotheses, then to synthesize the results through use of these hypotheses. The interpretation to be given the hypotheses is not that they have been verified or confirmed, but rather that on the basis of available evidence they appear to be promising leads for further systematic research.

In the perspective of assertions by the writer in earlier publications which were extensively concerned with the development of a research strategy in the domain of teacher behavior and effectiveness, the present statements may be viewed as strategy modifications and extensions, occurring in the latter stages of strategy development, but still as essentially pre-theoretical. The assumptions and hypotheses utilized lack both the degree of confirmation and the precision required of propositions which can stand as materially true synthetic statements in a formal postulate system.

Hypothesis Generation

In the present study three main categories or sets of variables have been recognized. (1) Those associated with the institutional context within which the teacher teaches. (2) Those associated with work tasks of the teacher and (3) Those associated with the teacher as a person or with the "personal context" generated by the teacher. Collaterally, the principal problems to be dealt with are 1. to relate variables within these sets, 2. to relate variables between sets, and 3. to relate variables and sets within the study to variables, sets or theoretical statements essentially outside the study.

Relationships Between Variables Within Sets. Of the three sets of variables examined in the study those in the set labeled "institutional context" occupied a superordinate position in the sense that they controlled or moderated relationships between variables belonging both to this set and to other sets. Within the set of variables comprising the institutional context several sub-categories, groups or sub-sets may be distinguished.

One sub-set was comprised of social and economic variables, including the adjusted assessed valuation per resident pupil ADA for the school system, the degree of industrialization of the community and the general socio-economic class background of the children in the school system. Within the sample of systems used in the study, the socio-economic background of the patrons of the system varied inversely with adjusted assessed valuation and the collaterally associated degree of industrialization. In a broader sample of school systems, however, this inverse relationship will undoubtedly appear in connection with only two out of a minimal four cell arrangement. The current sample was dominated by high assessed valuation systems with many working class patrons, and by low assessed valuation systems with many middle class patrons. Systems with high assessed valuation and many middle and upper middle class patrons (e.g. very wealthy suburban systems) and systems with low assessed valuation and many working class patrons (exemplified by system C in the present sample) clearly exist. Utilization of data from such systems in future studies should bring into sharper focus how the wealth of a system and the predominate class background of its patrons bear on such matters as supervision and differential valuation of teacher characteristics.

Generally, the variables appearing in this first sub-set may be conceptualized as providing the broad socio-economic conditions within which school systems must function. As these conditions vary the characteristics of school systems may be anticipated also to vary.

The second sub-set of variables within the set "institutional context" encompassed five variables, which may be identified as characteristics of school systems. Three of the five were observed variables and two were inferred from variations in variables outside this sub-set. The three observed variables were organization of supervisory services, amount of supervision given to beginning teachers and leniency errors in appraising beginning teachers. The two inferred variables were emphasis on teacher characteristics during supervision and emphasis on teacher characteristics in making supervisory appraisals. These latter two variables are clearly elements of the sub-set under consideration; however, because they were not independently defined, i.e. reduced to observable variables in this sub-set, they will be discussed more fully in conjunction with the sub-set from which they were inferred.

Of the observed variables in this second sub-set, the organization of supervisory services and leniency errors in appraising beginning teachers varied with the assessed valuation of the system. High assessed valuation systems provided a relatively high ratio of supervisors to beginning teachers with many contacts between supervisors and teachers, while low assessed systems provided a relatively low ratio of supervisors to beginning teachers, with few contacts between supervisors and teachers. Low assessed valuation systems tended to pass the supervisory function to the principals, while high assessed systems split this function among specialized supervisors and principals.

The association between leniency errors in appraising beginning teacher success and assessed valuation has several dimensions. First, principals in low assessed valuation systems were not only more lenient than principals in high assessed systems, but within each assessed valuation group principals were more lenient with primary than with intermediate teachers. Second, supervisors and superintendent in high assessed systems showed no leniency errors in their appraisals, while supervisors and superintendents in low assessed systems, like their principals, did show leniency errors.

This arrangement of leniency errors must first of all be observed to have occurred in relation to a group of beginning teachers who were relatively homogeneous across systems and across grade levels on the characteristics measured. While significant differences on MTT scores between systems and between grade levels were present in post-test performance, in pre-test performance on this instrument and on the other instruments used, the groups were quite similar. Leniency errors, therefore, cannot reasonably be attributed to differences between teachers by system and grade level, rather, these errors are to be accounted for primarily by means of variables operating within the institutional context.

The fact that principals in both high assessed and low assessed systems were more lenient with primary than with intermeidate teachers indicates that one source of leniency errors lies in the principals themselves. Given the fact that virtually all of the 73 principals involved in the study were males, and given the fact that virtually no males are to be found teaching the primary grades, the hypothesis emerges that as a group principals have less well elaborated criteria on which to gage the teaching of primary teachers than they have the teaching of intermediate teachers. The result is a tendency to overrate primary teachers.

While this hypothesis might be shown to be tenable upon subsequent research, it is clearly quite inadequate to account for the finding that supervisors and superintendents, as well as principals, in low assessed systems were more lenient in their appraisals of beginning teachers than were these persons in high assessed systems, and that leniency errors between grade levels in high assessed valuation systems were removed when only supervisory appraisals were used. Two hypotheses might be entertained to account for these latter findings.

First, the way in which supervisory services were organized in the high assessed systems intrinsically permitted a better sample of teacher behavior in these systems. As shown earlier, the mean amount of supervision per teacher given by principals in high and low assessed valuation systems did not differ; however, because of the greater prevalence of supervisors in high assessed systems, each beginning teacher received a bonus in supervisory contacts from this source. While this bonus had other apparent effects, to be discussed later, the effect of interest in the present context is the increase in the sample of teacher behavior gained by the school system through having had a supervisor as well as a principal in the classroom of the beginning teacher, thus potentially increasing the accuracy of the appraisals made.

The first of the two hypothesis to account for differences in leniency errors between systems, then, is that a school organization which provides supervisory services in addition to the principal, also intrinsically provides for a better total sample of teacher behavior and thereby increases the probability of a relatively unbiased rating.

The second hypothesis to account for differences in leniency errors between systems is derived centrally from the outcomes of the ratings themselves. Low assessed valuation systems were shown to place emphasis on the personal-social characteristics of the teacher, with the warm, friendly, imaginative teacher who had a favorable attitude toward pupils being favored as the more successful. These systems thus seem "person oriented." High assessed systems, on the other hand, favored task performance, especially in arithmetic, and, with respect to the types of problems beginning teachers had, were concerned with discipline and classroom organization problems, which were in turn associated with low scores on Ryans' businesslike, responsible Y scale.

These differences between systems suggest the possibility that low assessed systems operated on a more "person oriented" criterion than high assessed systems, and were willing to give teachers the benefit of doubt in appraising there over-all success.

While both of these hypothesis account for the differences in leniency errors between systems about equally well, the first seems to account somewhat better for the reduction in leniency errors in supervisory appraisals in high assessed valuation systems compared to those of principals. Supervisors had a better total sampling base of beginning teacher behavior than principals since they saw a wide range of beginning teachers. Principals, on the other hand, were restricted to a single building and frequently to a single beginning teacher, thus, the supervisor in each high assessed systems was in a better position to appraise each beginning teacher against the full range of such behavior among both primary and intermediate teachers than was any other person in the system. Perhaps it was this fact which lead to a better distribution of appraisals from supervisors. It might be noted in this respect that in many, although not all, low assessed systems the appraisal of the performance of each teacher by the superintendent must have been predicated primarily upon the recommendation of the principal since the principal was the only person in the system directly responsible for supervision.

These latter statements are not intended to eliminate the second hypothesis; however, the line of associations within the total set of variables (those describing the institutional context) under consideration appear to the author to be as follows: that the assessed valuation of the system is at least associated with, if not a controlling factor in, the presence of an elementary supervisor or supervisory staff; the presence of a supervisor or supervisory staff is in turn associated with (1) a greater total amount of supervision for the beginning teacher, and (2) less leniency error in appraising beginning teachers, which in turn may be hypothesized to be a function of a better sampling of teacher behavior

The second set of variables examined in the study comprised the performance of teaching tasks. Within the set two subsets were used, tasks in teaching arithmetic and tasks in teaching reading. Each of these subsets was in turn identified with four variables, representing task scores, and two additional variables representing the sum of task scores on each instrument, labeled MTT and TTR scores.

Under factor analysis the MTT and the TTR loaded on orthogonal factors when the sum of task scores on each of these instruments was left in the underlying correlation matrix along with the individual task scores. Removal of the total scores from the matrix, however, produced a set of factors which showed moderate relationships between certain MTT task and certain TTR tasks, primarily among intermediate teachers. The pattern of factor loadings on the factors extracted for the group of intermediate teachers suggested three distinctly interpretable factors. First, a factor identified with ability to form concepts or categories.(factor 2, Intermediate teachers, Table 41). This ability is presumed to underlie general skill in diagnosing pupil difficulties in learning, but in the present sample it appeared most clearly in connection with tasks in teaching arithmetic and mucn less clearly in connection with tasks in teaching reading. Thus, this factor cannot be interpreted as a pure "category formation" factor, since it appears to have distinct arithmetic components.

In terms of the three main work tasks of the teacher, i.e. pupil task setting, the appraisal of pupil responses against criterial responses and "instruction", this factor may be viewed as a major component of the appraisal of pupil responses against criterial responses. Viewing the factor in this context, the pattern of factor loadings on it seems highly consistent. The appraisal of pupil responses against criterial responses neccessarily requires a knowledge of what the criterial responses are like, which entails a knowledge of the content area in which the responses are made, since "knowledge" of a content domain by definition includes the availability of correct or criterial responses in that domain by the person holding the knowledge. Hence a factor germane to this domain would be expected to have a content component, which in the present case consists largely of arithmetic.

When the appraisal of pupil responses occurs in a context of diagnosis, rather than a context of evaluation of status, however, another dimension is clearly added to the task, namely, the diagnostician must be able to collect and group responses into categories such that an inference as to the nature of the difficulties underlying erroneous responses is possible, i.e., an induction or act in concept formation must be made. Thus a second component of the performance of this work task emerges, viz., a concept formation component.

The absence of an identical factor among primary teachers (although factor 5, Primary Teachers, Table 41 is similar) is probably predicated on the fact that the content of the arithmetic tasks is largely foreign to primary teachers, making it difficult for this group to supply the criterial responses antecedent to category formation, thus precluding the emergence of a factor with concept formation components in this group.

The second factor bearing on task performance extracted for the group of intermediate teachers (factor 3, Intermediate Teachers, Table 41) was identified with ability to program or order materials. Again this factor contains a definite arithmetic component as well as a reading component, but it cuts somewhat better across content areas than did the first factor. In terms of the work tasks of the teacher, this second factor may be viewed as a component of pupil task setting, containing elements which cover relating pupil tasks to expected outcomes, relating sets of pupil tasks to each other, and the sequencing of pupil tasks, all of which are interpreted to involve ability to bring order to materials.

Like the first factor, the second factor does not emerge with any clarity among primary teachers. Again, its failure to emerge appears to be traceable to the irrelevance of the content of the arithmetic tasks for primary teachers.

The third factor (factor 5, Intermediate Teachers, Table 41) bearing on task performance among intermediate teachers was identified primarily with teaching word attack skills in reading. This factor was similar to factor 2 among primary teachers, and appears to be tied almost wholly to the content or skill domain of reading. In the context of the work tasks of the teacher, this factor is ambiguous since a knowledge of specific content to some degree underlies all the work tasks of the teacher, and there is nothing in the factor loadings for this factor to indicate that it might be more closely attached to one type of work task than another.

With respect to the lines of association within the set of variables constituting task performance, the evidence available suggests very weak associations between total task performance in arithmetic and reading, i.e. between content areas, but moderate associations between the variables within these subsets which involve similar cognitive operations. It might be noted in this respect that better order could undoubtedly be brought to the variables in this set if measures saturated with arithmetic achievement, reading achievement, concept formation ability and ability to order materials respectively were placed in the matrix together with the individual task scores. If the components which appear to exist in the tasks truly exist in them, they should show reasonable loadings on factors known to be saturated with these components.

The third set of variables examined in the study were those identified with the teacher as a person, or the personal context generated by the teacher. Nine variables from the TCS comprised this set. As indicated via factor analysis, this set of variables was extremely homogeneous, with most of them loading heavily on a factor describable as a favorable personal-social orientation, although factors Y, B and I showed smaller loadings than did the other variables. The loadings of these latter variables on the personal-social factor were sufficiently modest to suggest that at least some of the variance of them might load on one or more other factors, although no explicit factor for the group or any single one of them could be identified under the orthogonal rotation of factors employed.

Relationships Between Sets of Variables. The relationships between the three sets of variables in the study were predominantly of the moderated independent-dependent or moderated predictor-criterion type. Thus, there were virtually no "straight forward" results; that is, results in which one variable was directly associated with another without having to take other variables into account.

The principal moderating variables appeared in the set labeled "institutional context." Of the variables in this set three, adjusted assessed valuation, supervision beyond that given by the principal and the ratio of lower to middle class schools varied together in such a was as to permit the identification of the two patterns among school systems: pattern A, characterized by high assessed valuation, substantial supervision beyond that given by the principal and approximately equal proportions of schools with working class and middle class children; and pattern B, characterized by low assessed valuation, little supervision beyond that given by the principal, and approximately two schools with middle class children for each school with working class children. These two patterns were subsequently treated as the two components of a dichotomous variable, and it is this variable which, along with a dichotomy of grade level taught, primary or intermediate, performed the principal moderating functions in the study.

The general line of moderated associations involving pattern A and pattern B systems began in the examination of the amount of supervision given by principals, as one independent variable and grade level as the other, with differences between pre-test and post-test scores on the MTT and the TTR providing the primary dependent variable and pooled pre-test, post-test scores on these devices the secondary dependent variable. The results of the analyses constituting the examination indicated that both system type and grade level were moderating variables in the relationships between amount of supervision and changes in task performance, but that the nature of the relationship was slightly different for the MTT than for the TTR.

Changes in MTT performance were directly associated with the amount of supervision received by intermediate teachers in pattern A systems such that those who received the greater supervision (> 1 hour per month) increased their MTT scores significantly more than those who received the lesser supervision. This effect, however, did not hold either for primary teachers in pattern A systems, or for intermediate or primary teachers in pattern B systems, thus it seems clear that the effects of supervision were moderated both by grade level and system type.

In order to account for these results it is first necessary to recognize that the MTT is grade-level-bound instrument. Typically primary teachers do not perform the type of tasks contained in this device in the classroom, hence no significant gain in performance would be expected with teaching experience. If primary teachers had gained significantly, confidence in the instrument would be weakened, since there are no hypotheses in the construct underlying the test which predicts that irrelevant experiences lead to significant changes in performance. Quite the contrary, as the author has elsewhere pointed out, (12) the prediction is that irrelevant experiences will not lead to performance change. In any event, the grade level specificity of the MTT undoubtedly accounts for performance differences between primary and intermediate teachers.

In order to account for differences in the affect of amount of supervision on changes in the MTT performance of intermediate teachers by system type, two hypotheses may be entertained. First, it may be that principals in pattern B systems place little emphasis on teacher task performance in arithmetic. This relatively slight emphasis on arithmetic teaching might be traceable in turn to the composition of the school population, which in these systems is predominantly middle class. Middle class children would be expected to have less difficulty learning the skill subjects than working class children, hence the need to emphasize a subject such as arithmetic would be comparatively reduced in predominantly middle class schools. Moreover, emphasis on arithmetic may go partially by default in these systems. That is, it might be a matter of emphasizing other aspects of teaching among beginning teachers, such as getting along with pupils and parents and keeping an attractive classroom, rather than emphasizing arithmetic.

Another hypothesis is that the greater total amount of supervision received by teachers in pattern A systems as compared to teachers in pattern B systems is a factor in changes in arithmetic performance. Teachers in pattern A systems not only had the benefit of supervision from principals, but in almost all instances the benefit of considerable help from a specialized supervisor or supervisory staff explicitly concerned with instruction.

Before further ordering these two hypotheses to account for the differences in the amount of change on the MTT between intermediate teachers in pattern A and B systems, the extent to which either can account for performance change on the TTR needs also to be examined.

On the TTR both primary and intermediate teachers in both pattern A and pattern B systems showed significant score increases. Among intermediate teachers in pattern A systems, however, there were differential increases such that those receiving more than one hour per month supervision from the principal gained significantly more than those who received less than one hour per month. On the grounds that a group of eight teachers who taught elementary school, but did not teach reading, failed to show any gain in TTR task performance the general significant gain among primary and intermediate teachers in pattern A and B systems was attributed to experience in teaching reading. This interpretation leaves open, however, the question why intermediate teachers who received much supervision in pattern A systems gained more than those who received little supervision.

Both of the hypotheses entertained with respect to the differential gains of intermediate teachers in pattern A systems on the MTT appear to account equally well for the similar effect obtained with the TTR. The first hypothesis, however, has an implication which the second does not. If principals in pattern B systems did not emphasize arithmetic or reading in their supervision, one might reasonably expect this fact to be reflected in differences in the hours per week class schedules for these subjects in pattern A and B systems. Yet, such differences did not appear. One must recognize, on the other hand, that hours per week class schedule is not the only index of the emphasis a given subject receives in a school. Nonetheless the absence of differences in hours per week class schedule between types of systems clearly does not strengthen the hypothesis.

Of the two hypotheses, the second is given preferential position by the author on two grounds. First, since the effects in question are directly attributable to amount of supervision by the principal, it seems logical to assume that adding supervision by a supervisor simply strengthens the effect. To account for the effect by introducing "emphasis" as an aspect of supervision seems somewhat less parsimonius that assuming that added supervision has cumulative effects on performance. Second, the assumption that additional supervision has a cumulative effect on performance fits well with reinforcement theory in the sense that a greater number of reinforcements or contingencies can be provided through the increased contacts of supervisory personnel with the teachers.

Acceptance of the second hypothesis as superordinate among the two which were ventured, and especially acceptance on the grounds that the second fits better with reinforcement theory, in turn implies an acceptance of supervisory personnel as one locus of reinforcement in the school system. More generally, it implies that one source of reinforcement for teachers lies in the authority structure of the work organization to which they belong. It is clear, however, that there is another locus of reinforcement for teachers. TTR performance changed significantly for all primary teachers and for intermediate teachers in pattern B systems wholly independent of amount of supervision received. Among intermediate teachers in pattern A systems significant effects for gains independent of an interaction between gains and supervision were also obtained. These results imply that there is another locus of reinforcement for teachers in their work. While the exact source cannot be directly inferred from present results, it is assumed to lie in the pupils taught.

In observing differences in the pattern of changes on the MTT and TTR, it appears to be the case that changes in MTT performance are wholly associated with the authority structure of the school as the source of reinforcement, insofar as no changes in performance occurred independent of the amount of supervision received. Changes in TTR performance on the other hand appear to be primarily associated with another locus of reinforcement, inferred to lie among the pupils taught, and secondarily associated with reinforcements having their locus in the authority structure of the school.

These differences in the loci of reinforcement inferred to exist between the MTT and the TTR cannot be fully clarified on the basis of the data available in the study, however, one hypothesis to account for them might be ventured. As a group elementary school teachers are recognized as being threatened by mathematics, including arithmetic (4, 5)*. Assuming this observation is true, it seems distinctly possible that many elementary teachers have strong avoidance responses toward this subject, which preclude effective changes in performance via reinforcement from students. That is, the "set" of teacher is to avoid the field, but since the authority structure of the school will not permit complete avoidance, the teacher may respond minimally to the field and leave it as rapidly as possible, ignoring the potentially reinforcing stimuli from pupils. Thus, in order for performance change to occur in arithmetic, the introduction of reinforcement, or the joint introduction of threat of deprivation (viz., implicit threat to deprive a beginner of his position) and positive reinforcement via the supervisor may be required.

---

*According to Duttons results(4) 74 percent of the preparatory elementary teachers beginning his classes showed strongly emotional unfavorable attitudes toward arithmetic.

In connection with this hypothesis it may be observed that low scores on both the MTT and TTR were associated, although rather weakly (p. $<$.10) with amount of supervision received by intermediate teachers in pattern A systems. While these associations must be regarded with caution, they nonetheless suggest that increased supervision is associated with poorly developed responses, which in arithmetic and possibly in reading, might be anticipated to also be associated with avoidance responses toward these subjects among certain teachers.

A question may be raised at this point concerning why supervisory personnel in type A systems might bother either to reinforce or to threaten deprivation in connection with arithmetic and reading among intermediate teachers, while principals in pattern B systems apparently would not bother with such matters. To answer this question, even tentatively, it seems necessary to appeal to the hypothesis that there are underlying differences in the emphasis on skill subjects between the two types of systems, probably predicated on the underlying differences in the socio-economic class composition of pupils in them. Thus, in order to account for the effects of amount of supervision on performance change, both of the hypotheses initially ventured must, at least for the present, be retained.

Before terminating discussion on the points immediately above, it is of some interest to note that the significant changes in MTT scores among intermediate teachers in pattern B systems who participated in in-service work related to teaching arithmetic are fully consonant with the hypotheses made. MTT performance change was wholly traceable to sources of learning outside of the classroom, and in the instance of the teachers in question, predominantly outside of the pattern B school systems.

The further operation of system type as a moderating variable were observed in a complex of results linking various variables to the amount of supervision received and the the types of problems beginning teachers experienced. In pattern B systems as a whole no variables were found to be associated at $p = .05$ either with the amount of supervision received or with the kinds of teaching problems teachers had as perceived by principals and superintendents. Moreover there was no association between the amount of supervision received and having teaching problems in these systems. Of these results, the disassociation of the amount of supervision from the problems experienced by the teacher is most curious. The absence of a relationship makes it appear that the supervision of many principals in these systems, although certainly not all, was organized around some concept or principle unrelated to the problems the teacher experienced.

In pattern A systems, the situation was quite different. First, the presence of discipline, organizational and control problems was significantly associated ($p < .005$) with the amount of supervision received. Second, TCS Scales B and Z were significantly associated with the amount of supervision received by intermediate teachers while MTT and TTR pre-test scores were also weakly associated with this criterion. And third, TCS Scales B, Y and X significantly differentiated between pooled primary and intermediate teachers who experienced no problems and those with discipline, organization or control problems, while Scale B also discriminated among those with no problems, those with miscellaneous problems and those with discipline and related problems.

The line of associations appearing in these results, in conjunction with results previously discussed, suggest three interrelated hypotheses relevant to pattern A systems. First, that discipline, control and organizational problems among beginning teachers are at least in part attributable to the conjunction in a teacher of a subject centered, traditional viewpoint, but a relatively poor command of skills in teaching arithmetic and reading, and difficulty in maintaining organized, business-like classroom behavior. Second, that the presence of these problems attracts increased supervision, which among intermediate teachers leads to increased skills in teaching arithmetic and reading, i.e. to performance change. Hence, third, that there is an indirect but inferrentially firm relationship between certain personal-social variables, or certain aspects of the personally generated context of teaching and performance change, such that problems predicated on difficulties arising from personal-social variables lead to increased supervision, which leads to performance change.

The line of associations on which these hypotheses are based did not hold for teachers in pattern B systems. The fact that they did not indicates a need for further hypotheses. Before proceeding to these hypotheses, however, the further effects of system type and grade level are moderating variables require examination.

One of the points at which the moderating effects of system type and to a lesser extent grade level was most apparent was in the differential relationship of the TCS variables and task performance variables to supervisory appraisals. In pattern A systems MTT pre-test performance among intermediate teachers was the only significant predictor of supervisory appraisals of "success", while in pattern B systems, the TCS variables loading heavily on Factor 1, Table 41, and especially scales X, Z, R and R_1 were positive predictors of success on the criterion of supervisory appraisals, with the MTT being a significant negative predictor. In addition, the association of the TCS scales with success in pattern B systems was shown to be identified primarily with teaching in middle class schools, but clearly not identified with teaching in working class schools in such systems, nor with teaching in either middle class or working class schools in pattern A systems.

These results, together with those related to the problems experienced by beginning teachers, amount of supervision received, and changes in task performance, all moderated by system type, suggest that there were basic differences in outlook between pattern A and pattern B systems, with some differences occurring within system types according to grade level. First, pattern A systems continually gave indications of strong "task orientation," or a heavy emphasis on task performance among intermediate teachers. The indicators are performance change on the MTT and TTR with high supervision, associations, albeit weak, between amount of supervision and pre-test performance on the MTT and TTR, and very strong associations between supervisory appraisals indicating success and pre-test performance on the MTT. The failure of the TTR to share the same level of association as the MTT with respect to supervisory appraisals unfortunately suggests that this instrument is weak and needs substantial working over in order to increase its validity, which was not well established when the study began. In pattern A systems, however, there was little indication of a heavy emphasis on task performance among primary teachers. Whether the failure to show task emphasis among these teachers is a characteristic of these systems or a statement about the relevance of the MTT and TTR for primary teachers is, of course, a focal but not fully answerable question in the present study.

Compared to pattern A systems, pattern B systems gave no indications of "task orientation." MTT performance among intermediate teachers in these systems changed only under influences from outside the system. Moreover, the pre-test MTT performance of these teachers was negatively related to success on the criterion of supervisory appraisals. Finally the increases in TTR scores among both primary and intermediate teachers in these systems occurred independent of any identifiable source of reinforcement in the authority structure of the system. Thus pattern B systems may be characterized as "low task oriented" systems, while pattern A systems may be characterized as "high task oriented" systems.

A second difference in outlook between pattern A and B systems is reflected in the differential relationships of the TCS scales to success on the criterion of supervisory appraisals. The association of scales X, Z, R, R_1 with success in pattern B systems, and the additional association of scales I and S with this criterion among teachers in middle class schools in predominantly middle class systems, is a strong indicator that pattern B systems emphasized personal-social behaviors. In pattern A systems, on the other hand, the TCS scales which load heavily on factor 1 were not associated with success. These systems may consequently be characterized as low in personal-social orientation. In general, pattern A systems may be viewed as high task oriented, low personal-social oriented; while pattern B systems are low task oriented, high personal-social oriented.

There were indications, however, that some personal social variables, especially scales B and Y, which do not load heavily on factor 1, were relevant to certain aspects of success in pattern A systems, specifically to the occurrence or non-occurrence of teaching difficulties. At the same time, it appears to be the case that the association of these two variables with problems experienced are not indicators of the orientation of the system. The orientation of pattern A systems is indicated by the fact that discipline and organization difficulties are closely associated with both amount of supervision received and supervisory appraisals of success. These systems thus especially emphasized freedom from discipline and organization problems in the criteria of success. Yet, neither scale B nor scale Y, nor any TCS scale was directly associated with supervisory appraisals of success in these systems. The situation appears to be one in which the personal-social variables which underlie teaching difficulties or the absence thereof are not in themselves either valued or disvalued in pattern A systems. Rather, the associations these variables hold to teaching difficulties appear to reflect intrinsic relationships in the sense that a subject centered viewpoint but the absence of businesslike, systematic behavior leads to discipline and organization difficulties, at least in systems having a large proportion of pupils from working class homes. To put the matter another way, supervisory personnel in pattern A systems were not saying "we value teachers who are permissive" or "we value teachers who are systematic," but they were saying "we value teachers who are free from discipline and organization problems." If it is the case that teachers who are free from discipline problems in pattern A systems are child centered but systematic, that is a matter of fact, not a matter of value or preference.

The general interpretation to be made of the differences in orientation between pattern A and pattern B systems is heavily dependent upon experimental work described by Robert F. Bales in examining task roles and social roles in small problem solving groups (1). The groups consisted of four or five persons who were to solve a problem or task. Behavior during task performance was recorded and resolved via interaction analysis; subsequently, each member of the group rated the other members as to likability and as to contribution to the task, or "ideas." Out of this type of situation three factors each of which characterize leadership roles were identified. The factors are uncorrelated (orthogonal), but not mutually exclusive, hence it is possible that an individual may play all three roles, although it is not likely that he will do so. It is more likely that different members of a group will assume different roles, thus complementing each other. The first role may be characterized as that of "task specialist." The person in this role supplies many ideas relevant to the task and is "active" in the sense of responding or talking considerably in the group. The second role may be characterized as that of the "social specialist." The person in this role specializes in social-emotional problems in the group, gives reinforcement or positive reactions, receives problem-solving attempts, but does not himself talk extensively. The third role is that of the "overactive deviant." The person in this role talks too much, attempts to dominate, but is relatively low on task ability.

Likability, as rated by other members of the group, varies with these three roles. The most likable person is that who plays the "social specialist" role, the least likable, the "overactive deviant." Likability is in turn linked to the amount of feedback given by a person in each role. A "task specialist" who receives many responses from others in proportion to the responses he initiates is thus much better liked than the task specialist who is "over active", as it were. To put the matter another way, the task specialist who deviates toward the role of the social specialist will be more likable than the task specialist who deviates toward the role of the overactive deviant.

Assuming that the classroom group is parallel with, although by no means identical to, the small problem solving group, and therefore that the roles described by Bales apply, a number of hypotheses may be generated through which a substantial portion of the results of the present study may be interpreted. First, it appears to be the case that pattern A systems emphasize or value the role of "task specialist" among intermediate teachers. Whether a similar role is emphasized among primary teachers in this type of system remains to be determined in future studies in which teaching tasks more nearly relevant to primary teachers are utilized. Second, pattern B systems appear to emphasize the role of "social specialist." This emphasis appears to run through both the primary and the intermediate grades in such systems. Third, the relevance of TCS scales B and Z to the amount of supervision received by intermediate teachers in pattern A systems, and the relevance of scales B, Y and X to the occurrence of teaching problems in these systems suggest a second, probably subordinate role for intermediate teachers, but possibly superordinate for primary teachers in these systems. Of these scales, B held the dominant position for both primary and intermediate teachers. Given the assumption that this scale is a valid measure of child centeredness, the hypothesis is suggested that the task specialist role in pattern A systems is most successful when coupled with child centeredness, which is interpreted to be correspondingly associated with positive reactions and high feedback, and with likability. In pattern A systems, Factor X (friendly, understanding) and Factor Z are probably aligned with Factor B in this role,* while Factor Y is probably aligned with the task specialist role. The task orientation of the task specialist suggests that persons in this role would be systematic, businesslike and orderly in their appraoch, hence relatively high on scale Y, which was the case for pattern A teachers who experienced the fewest problems.

---

*If the Ryans studies (8, pp 231-232), Scales X, Z and B are more closely associated than they are in the present study, and load heavily on a common factor.

Given this interpretation of differences in orientation between pattern A and pattern B systems, the question to be answered is whether differences in the relationships between the problems experienced by teachers and (1) amount of supervision, and (2) supervisory appraisals by system type can be resolved by means of the interpretation.

To account for these relationships, the following hypotheses may be entertained. In pattern A systems, task specialization may be seen as the superordinate or most valued role. Since this role is most valued, classroom problems which are inconsistent with this role, viz., discipline and control problems, become a locus of concern for supervisory personnel. Consequently, a second, subordinate role emerges as significant, viz., the role of social specialization. This role, however, is not valued in and of itselfin pattern A systems, rather, the role emerges because taking it is intrinsic to the avoidance of discipline and organization problems. As Bales indicates a person in the task specialization role who fails to give feedback and positive reactions generates hostility and is disliked. This interpretation would then account for the relationship between teaching problems in discipline and organization and the amount of supervision received in pattern A systems. It does not account fully, however, for the association between these teaching problems and supervisory appraisals of success, but the failure to find an association between the TCS variables and appraisals of success. The latter failure in association where one might be anticipated seems traceable to two sources: (1) the subordinate position of the social specialization role, a role intrinsic to successful teaching but not emphasized in pattern A systems, and (2) the fact that the supervisor can take the changes in teacher performance brought about by principals into account in making his over-all appraisal. In a number of cases supervisors said "Teacher X had difficulties at the beginning of the year, but she has responded beautifully and is now a superior teacher." It might be noted that this same factor appears to underlie the failure of amount of supervision received to correlate with supervisory appraisals of success in spite of the fact that the problems experienced by a teacher were significantly associated with each of these criteria.

In contradistinction to pattern A systems, the social specialist role in pattern B systems may be hypothesized to be superordinate while the task specialist role seems not merely subordinate but negatively valued or disvalued. This arrangement of relationships is remarkably similar to that found by Bales, in that the task specialist appears to be the person less well liked, if he is overactive. Correspondingly it appears that pattern B systems heavily stressed likability among their teachers, which in turn would account for the emergence of scales X, (warm, friendly, understanding), R (favorable attitude toward pupils) and R_1 (favorable attitude toward democratic pupil practices) and S (emotional adjustment) as significant correlates of supervisory appraisals of success. Collaterally, the negative association betweeen MTT performance and supervisory appraisals may be interpreted as reflecting a negative association between task specialization, especially when coupled with overactive responding or "dominance", and likability.

That over-active responding must be coupled with task specialization in order for a negative association to appear is a critical point. As the loadings of the MTT and TTR tasks on factor 1 in Table 41 indicate, task performance and the personal-social factor associated with social specialization are orthogonal, a finding consistent with those of Bales. It is only when coupled with "dominance" that a negative association between task specialization and likability would be expected to occur.

The failure of the discipline and organizational problems experienced be teachers in pattern B systems to be associated with amount of supervision received, but to be positively associated with supervisory appraisals may be interpreted as follows. The low emphasis on task specialization in pattern B systems would not create the inconsistency between task specialization and discipline problems noted in pattern A systems. Principals in pattern B systems would therefore not show a focal concern with such problems, and they do not in fact express the problems of teachers in terms of discipline and organization, but in terms of management. The low level of concern with such problems would consequently not lead to greater supervision, hence amount of supervision received and problems experienced by teachers would be uncorrelated in pattern B systems.

The association between problems experienced and supervisory appraisals of success in pattern B systems appears to rest on somewhat different grounds. First, examination of Table 35a indicates that this association rests on the disproportional absence of problems among teachers appraised as above average-superior, rather than on the disproportional presence of problems among teachers rated average below-average. To put the matter another way, unusually large proportions of teachers who have problems were rated above-average superior by supervisors. Nearly two thirds of the teachers with miscellaneous problems and nearly one-half the teachers with management problems were rated above average-superior. These proportions, it may be noted, are quite unlike those in pattern A systems as shown in Table 35. Second, it was in pattern B systems where the leniency errors of supervisory personnel were most pronounced, hence, in general, a disproportional number of teachers were appraised as above average-superior, although such leniency errors were not present in the amount of supervision received. These results suggest that in making appraisals of success, supervisory personnel in pattern B systems focussed on teacher characteristics which are linked to, but not closely associated with teaching problems. Too many teachers who had teaching problems were appraised as "successful" for close association to be present. Only a part, actually about 42% of the total group of teachers who had teaching problems were appraised as average-below average. The average-below average teachers moreover, were shown to be lower on those characteristics associated with social specialization than were teachers appraised as above average-superior.

This state of affairs suggests that the association between teaching problems and supervisory appraisals is a secondary rather than a primary association. That is, that the primary association is between appraisals of success and personal social variables, i.e. variables linked to social specialization. These variables are in certain instances also relevant to the occurrence of teaching problems, but compared to the total number of teaching problems occurring, the number associated specifically with variables in the personal social area is small. Hence, personal social variables would not be correlated with the occurrence of teaching difficulties as a general rule, and indeed, there were no such associations in pattern B systems.

This interpretation, together with the observed fact that the problems experienced by beginning teachers in pattern B systems were not related to the amount of supervision received from principals, implies that supervisory personnel in pattern B systems did not organize either the amount of supervision given or supervisory appraisals of success around the problems experienced by the teacher. A locus or loci around which pattern B principals might have organized the amount of supervision received (or "needed") by a beginning teacher could not be identified, while the locus around which supervisory appraisals were organized appeared to be the personal-social characteristics of the teachers, especially those which may be inferred to be associated with Bales' factor of social specialization, hence, likability.

Shifting at this point to the affects of teacher mobility on the data and interpretations under consideration, two points may be observed. First, among non-mobile teachers differences in emphasis between pattern A and pattern B systems with respect to supervisory appraisals remained. Pattern A systems continued to show a task orientation and pattern B systems a personal-social orientation. Second, because of variations among mobile teachers certain changes in score level occurred among non-mobile teachers. In both pattern A and pattern B systems, the less businesslike (scale Y) intermediate teachers were mobile, hence the total group of non-mobile intermediate teachers became more businesslike by virtue of differential selection. At the same time, the less friendly, understanding (scale X) primary teachers left pattern A systems, while the slightly more traditional teachers (scale B) left the primary grades in pattern B systems, but the slightly more permissive intermediate teachers left pattern A systems, resulting in significant differences in permissiveness between pattern A and B systems among non-mobile teachers. Differential mobility thus had a tendency to sharpen differences between pattern A and B systems, with the pattern A systems having a somewhat more traditional and more task oriented group of intermediate teachers by the end of the study than did pattern B systems, where mobility resulted in a general increase in permissiveness (Scale B) scores. While the differences between systems at the end of the study was certainly not marked or striking, the trend was toward greater differentiation following the lines of the dominant emphais in the two types of systems: toward greater task specialization in pattern A systems, and toward greater social specialization in pattern B systems, with businesslike behavior gnerally increasing in both types of systems, especially among intermediate teachers.

The above interpretations of the differences in emphasis between system types leaves essentially unanswered why such differences occur. The evidence available in the present study suggests that the answer lies in the fact that the personal-social emphasis of pattern B systems is directly centered in middle class schools in predominantly middle class school systems. Thus the emphasis on the personal social qualities of the teacher or a social specialization role appears to be aligned primarily with middle class values. The emphasis on task performance or task specialization in pattern A systems is correspondingly inferred to be more nearly aligned with working class or perhaps upper lower and lower middle class values, since the latter two classes are those primarily represented in the term "working class" as used in this study.

This interpretation is to a considerable degree reinforced by James Coleman's findings reported in The Adolescent Society (4). As Coleman shows, the values of mid western adolescents in such middle class suburbs as "Executive Heights" are solidly centered in "personality." That is, the personal attributes most valued by the middle class adolescent in a middle class suburb are those in the personal-social domain. "Scholarship" on the other hand, is not especially valued. In Coleman's industrialized "Mid city" in central Illinois, on the other hand, "scholarship" is valued to a substantial degree, while "personality" is relatively less valued although not disvalued.

Recognizing at the outset that adolescents in secondary schools are quite a different population than beginning elementary teachers, it nonetheless remains that each of these groups have a school system as the locus of their activities, and that both groups are immersed in, or at least subject to pressure from the communities of which they are a part. Without doubt the school system experiences almost exactly the same value pressure from a community that adolescents experience, hence the values absorbed and reflected by the one should be similar to those absorbed and reflected by the other. Indeed, it is not unthinkable that the values expressed by adolescents are learned in part through the school system, which mediates the values of the community not only for teachers, but also for pupils.

Synthesis of Results

Results may be synthesized at different levels of generality. In the preceding section the focus was on interpreting and relating results in some detail. The purpose of this section is to view the results somewhat more generally, and to link them to the concerns expressed in the opening chapter.

In viewing teaching as an occupation, the results very clearly suggest that a tripartite division of the occupation into its institutional setting or context, its work tasks, and the personal characteristics of the practitioner or the personal context generated by the practitioner, is at least highly advantageous, if not necessary. By viewing the occupation as having three major components at the outset of the study, it was possible to show that both the success of a teacher and the way in which teacher performance changes in the early years of teaching experience depends on the type of institutional context within which the teacher finds himself. Statements about teaching "success" were clearly demonstrated to be relational statements, that is, statements which express a relationship between the personal and task performance characteristics of the teacher and the characteristics of the institutional context in which he teaches. Teachers with very similar characteristics but teaching in different types of settings may find themselves at roughly opposite ends of the success continuum as defined by supervisory appraisals.

Which particular teacher characteristics will be associated with success in a given institutional context appears to depend primarily on the socio-economic class composition of the students in this context. When a large proportion of the students (approximately 50%) are of working class origin, the work task performance of the teacher appears as a controlling factor in success, with systematic, businesslike behavior but a relatively permissive viewpoint probably also being contributing factors. On the other hand, when a very large proportion of the students (60% or more) are of middle class origin, the personal-social characteristics of the teacher appear as controlling factors in success, with positive emphasis on such characteristics as friendly, understanding behavior; stimulating, imaginative behavior; and favorable attitudes toward pupils and democratic pupil practices, but negative emphasis on task performance.

The differential association of personal characteristics and work task performance with success in different types of institutional contexts not only followed the conceptual distinction between work task performance and personal-social characteristics made at the beginning of the study, it also followed the factor composition of these components. Work task performance and personal-social variables were found to load on orthogonal factors, which may in turn be seen as making the results possible from an empirical viewpoint.

Viewing the results in question in a different perspective, both the factorial composition of major variables involved in the study and the detailed relationships among results shows marked parallelism with empirically derived factors discussed by Bales with respect to leadership roles in small problem-solving groups. The emphasis on task performance in systems with many working class students suggested that the role of "task specialist" held a superordinate position among the criteria employed by supervisory personnel in these systems, while the emphasis on personal-social relationships in systems with many middle class students suggested that the role of "social specialist" held a superordinate position among the criteria employed by supervisory personnel in these systems.

The relationship between the institutional context and changes in task performance appear to be only slightly different than those between the former variable and teacher success. Changes in task performance on the MTT and to a less extent on the TTR appear to depend, broadly speaking, on both school system wealth (adjusted assessed valuation) and the socio-economic class composition of students. School system wealth appears to enter as a variable primarily because of its association with extent of supervisory service beyond those provided by the principal, which in turn appears to be a major contributing factor in the association between amount of supervision received and changes in performance. The socio-economic class background of students also appears to contribute to this association on the same basis that it enters the association between task performance and teacher success, viz., it appears to be a determinant of whether task performance or personal-social characteristics are stressed. Changes in MTT performance occurred as a function of amount of supervision received from principals only when there were supervisory services beyond those provided by the principal and when there were many students of working class origin in the school system.

Changes in TTR performance did not follow exactly these lines. TTR performance increased for all groups of teachers, primarily as a function of the classroom teaching of reading, with a differential increase occurring among intermediate teachers as a function of amount of supervision in systems with extensive supervisory services beyond the principal and with many students of working class origin.

Interpreted within a broad conceptualization of reinforcement theory, the changes in MTT and TTR performance and differences in the variables to which these changes were attributable according to which test was used as the dependant variable, suggest the following possibilities. The locus of reinforcement leading to changes in MTT performance appears to lie in the authority structure of the school system, and specifically in the actions of supervisory personnel. Thus, the institutional context appears to hold the determinante of performance changes in arithmetic. The locus of reinforcement leading to changes in TTR performance, on the other hand, appears primarily to lie elsewhere, presumably among the students taught, although a secondary locus of reinforcement among supervisory personnel for intermediate teachers in systems with extensive supervisory services and many working class students is suggested by the data.

The failure of primary teachers to increase in MTT task performance, but their significant increases in TTR performance may be traceable in part to the apparent differences in the loci of reinforcements speculated to underlie performance change on these tests. The controlling factor in these differential changes, however, undoubtedly lies in the differential relevance of the MTT and TTR to what primary teachers actually do in the classroom. Primary teachers are focally teachers of reading, not of arithmetic, especially of arithmetic involving use of concepts in division and fractions, which are central on the MTT. Indeed, if primary teachers had gained on the MTT, the validity of this instrument as relevant to classroom teaching would be seriously in doubt.

In a different perspective, the failure of primary teachers, especially primary teachers in pattern A systems, to increase in MTT performance, points up the complexity of the factors bearing on work task performance. Some teachers, namely those in pattern B systems, did not increase in MTT performance for what appear to be social reasons, i.e. performance of tasks of the type contained in the MTT appears not to be valued as instrumental in the social context in which these teachers teach. In pattern A systems, on the other hand, there is reason to think that the performance of tasks of the type contained in the MTT is valued, but that particular sample of tasks in the MTT is relevant only to the universe of tasks performed by intermediate teachers. Thus, from the viewpoint of teaching in the primary grades, the MTT may be seen as sampling an irrelevant universe of tasks.

In this same light, the TTR may be seen as sampling what appears to be a relevant universe of tasks for both primary and intermediate teachers in terms of what these teachers actually do. All groups of teachers actually teaching reading showed significant score increases on the TTR, while teachers not teaching reading failed to show score increases. Yet, performance of these tasks appears to be quite unrelated to the values of the school system as expressed through supervisory appraisals of success. There was no group for whom the TTR predicted success. It thus appears that at some point in the construction of the TTR there was a critical failure in sampling. The tasks sample a relevant universe, but they fail to capture the portion of the universe of teaching tasks in reading which distinguishes between the more successful and the less successful teacher from the viewpoint of supervisory appraisals.

Examination of the factorial composition of the MTT and TTR suggests the possibility that the TTR does not sufficiently sample tasks involving the concept formation or diagnostic skills of the teacher, but over-samples achievement related tasks, especially those involving a knowledge of work attack skills. This assumption is based on the fact that those tasks which load most heavily on the concept formation factor (Factor 2, Table 41) have, in general, the best discriminating power on the criterion of supervisory appraisals of success.

Viewed in the perspective of previous research and writing by the author, the total results of the present study may be seen as forcing a shift away from the position that problem solving proficiency, or in more nearly operational terms, task performance, is the defining characteristic of successful teachers, toward the position that task performance is one of several teacher characteristics which will appear as a predictor of teacher success, depending on the institutional context within which the teacher teaches. More generally, this shift in position is away from the assumption that statements about teacher skill, success, proficiency or effectiveness are statements about some attribute or set of attributes of a teacher toward the assumption that statements about teacher skill, success, proficiency or effectiveness are statemetns about the relationship between the characteristics of a teacher and the institutional context within which he teaches.

A second shift in position involves the assumption, expressed in earlier publications, that the central reinforcers in teaching lie in the classroom. This assumption overlooks the power of the authority structure of the school and in terms of present results appears to mislocate a principal locus of reinforcement for teachers, viz., supervisory personnel. Moreover, the findings that changes in teacher behavior are traceable to the actions of supervisory personnel raises a serious question about how specific instrumental responses could be reinforced by supervisors or principals since they are not frequently present in the classroom. Even the most highly supervised teachers in the present study did not encounter the principal and/or supervisor more than one hour per week in toto, an insufficient period for much reinforcing of specific responses to occur. Clearly, supervisory personnel must be dealing with teacher behavior in units much broader than specific response units. Certainly there is much to suggest that viewing teaching behavior as being modified at the specific response level is inadequate, which raises serious questions about how a reinforcement theory predicated primarily on observations of changes in specific responding under reinforcement applies to changes in the behavior of the teacher. The resolution of this question, however, awaits further conceptualization and the results of future research studies.

References

1. Bales, Robert F., "Task Roles and Social Roles in Problem Solving Groups," in *Readings in Social Psychology*, E.E. Maccoby, T.M. Newcomb, and Eugene L. Hartley (Eds.), Holt, Rinehart and Winston, New York, 1958.

2. Burnette, Richard W., Wade, Eugene W., and Davis, Nancy B., *Elementary Grades Teaching Tasks in Reading*, Institute of Educational Research, Indiana University, Bloomington, Indiana, 1961.

3. Bradtmueller, Weldon G., *An Investigation of the Development of Problem Solving Skills Related to Teaching Reading During Student Teaching*, Doctoral dissertation, Indiana University, 1963.

4. Coleman, James A., *The Adolescent Society*, The Free Press of Glencoe, Glencoe, Illinois, 1961.

5. Dutton, Wilbur H., "Attitudes of Prospective Teachers Toward Arithmetic," *Elementary School Journal*, Vol. 52, No. 2, October , 1951, pp. 84-90.

6. Dutton, Wilbur H., "Measuring Attitudes Toward Arithmetic," *Elementary School Journal*, Vol. 55, No. 1, September, 1954, pp. 24-31.

7. Gross, Edward, *Work and Society*, Crowell, New York, 1958.

8. Ryans, David G., *Teacher Characteristics Schedule*, Form E 54, 1954.

9. Ryans, David G., *Characteristics of Teachers*, American Council on Education, Washington, D.C., 1960.

10. Smith, B. Othanel, and Meux, Milton, *A Study of the Logic of Teaching*, Bureau of Educational Research, College of Education, University of Illinois, Urbana, Illinois. (undated).

11. Turner, Richard L., *Intermediate Grades Mathematics Teaching Tasks*, Institute of Educational Research, Indiana University, Bloomington, Indiana, 1960.

12. Turner, Richard L., *Problem Solving Proficiency Among Elementary School Teachers* II. *Teachers of Arithmetic Grades 3-6*, Monograph of the Institute of Educational Research, Indiana University, Bloomington, Indiana, 1960.

13. Turner, Richard L., *Problem Solving Proficiency Among Elementary School Teachers IV. Further Investigations of Teachers of Arithmetic Grades 3-6*, Monograph of the Institute of Educational Research, Indiana University, 1961.

14. Turner, R.L., White, K.P., Quinn, E.D. and Smith, N.W., *Skill in Teaching Assessed on the Criterion of Problem Solving: Three Studies*, Bulletin of the School of Education, Indiana University, Vol. 39, No. 1., Bloomington, Indiana, January, 1963.

15. Wade, Eugene W., *Problem Solving Proficiency Among Elementary School Teachers III. Teachers of Reading Grades 2-5*, Monograph of the Institute of Educational Research, Indiana University, Bloomington, Indiana, 1961.